Examining the role of cohesin subunit Rad21 in malignancy and development

Bridget Mackie

A thesis submitted for the degree of Master of Science

University of Otago

Dunedin

New Zealand

February 2020

ABSTRACT

The cohesin complex is essential for cell survival, owing to its well-established roles in cell division, DNA repair and gene expression. Malignancy and developmental disorders termed the cohesinopathies can result when mutations are present in cohesin subunits, or in proteins that interact within cohesin.

In this thesis, I describe experiments using the *rad21*^{nz171} zebrafish mutant line, which carries a null allele for cohesin subunit Rad21, to address cohesin's role in driving malignancy and normal development. The thesis is in two sections.

The first section examines the functionality of a germline RAD21 variant that may be responsible for the development of myelodysplastic syndrome in a cohort of patients. RAD21 functionality can be determined through the use of a zebrafish *runx1 in situ* hybridisation bioassay, and this research found that the familial variant produces a functional cohesin subunit. A novel RAD21 variant was designed and assayed alongside the familial variant, and found to be non-functional.

The second section focuses on the role of cohesin in regulating normal embryogenesis. High-throughput RNA-sequencing was performed on the elongating tailbuds of *rad21*^{nz171} zebrafish embryos compared to wildtype siblings, to elucidate which genes and biological pathways cohesin is fundamental in regulating during zebrafish embryogenesis. The tailbud was chosen for examination as the posterior of the zebrafish develops from a bipotent stem cell population in this structure, and a lack of cohesin-regulated gene expression is predicted to affect cell fate decisions of the stem cells. An abundance of metadata was generated from the RNA-sequencing analyses, providing numerous avenues for further exploration. The spatial expression of four genes identified in these analyses were examined via *in situ* hybridisation, providing a broader picture of cohesin-deficient development.

We discovered that in the absence of cohesin-regulation, the fate of bipotent stem cells in the developing zebrafish tailbud skews towards neural ectoderm over mesodermal progenitors, and theorise that this shift is due to dysregulation of *wnt3a*, a key regulator of zebrafish axial elongation. A role for cohesin in regulating ribosome biogenesis, and thus protein translation, was identified in these analyses. The involvement of cohesin in assisting in protein translation implicates cohesin as regulating gene expression not only at the gene, but protein level also – a role which has not been well explored in the literature to date.

ACKNOWLEDGEMENTS

Firstly, I would like to give a tremendous thank you to my supervisor, Associate Professor Julia Horsfield, for welcoming me to the Chromosome Structure and Development Group and providing guidance and support throughout the duration of my Master's research. To my fellow lab members, thank you for the laughs, encouragement and patience you have taught me. A special mention goes to Judith, who never once rolled her eyes at any of my mountain of questions.

To all those outside of the lab group who offered expertise throughout this research. Gregory Giminez, Joel Tyndall and Ben Halliday – your computer savviness is admirable.

Lastly, to all of my loved ones. Your support means the world, I'm so lucky to have you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

Symbol	**Meaning**
+/+	Wildtype
+/-	Heterozygous
-/-	Homozygous mutant
°C	Degrees Celsius
3D	Three dimensional
µg	Microgram
µl	Microlitre
AA	Amino acid
AML	Acute myeloid leukaemia
BMP	Bone morphogenic protein
bp	Base pair
CdLS	Cornelia de Lange syndrome
CTCF	CCCTC binding factor
dpf	Days post fertilisation
DNA	Deoxyribonucleic acid
ESCO2	Establishment of cohesion 1 homolog
FGF	Fibroblast growth factor
hpf	Hours post fertilisation
ISH	*In situ* hybridisation
HM	Haematopoietic mesoderm
kb	Kilo-base pair
L	Litre
LB	Luria broth
MDS	Myelodysplastic syndrome
MO	Morpholino
mRNA	Messenger ribonucleic acid
NIPBL	Nipped-B like protein

nL	Nanolitre
NMP	Neuro-mesodermal progenitor cell
M	Molar
MDS	Myelodysplastic syndrome
mg	Milligram
mL	Millilitre
mM	Millimolar
PCA	Principal component analysis
PCR	Polymerase chain reaction
pg	Picogram
qPCR	Quantitative polymerase chain reaction
RB	Rohon Beard neurons
RBS	Roberts syndrome
RNA	Ribonucleic acid
Rpm	Revolutions per minute
SDM	Site directed mutagenesis
SMC	Structural maintenance of chromosome
TAD	Topologically associating domain
V	Volt
WT	Wild type

NOMENCLATURE

The table below lists the specific gene and protein nomenclature for zebrafish and human. All gene names are italicized, and protein names are not.

Species	**Gene**	**Protein**
Zebrafish	*rad21*	Rad21
Human	*RAD21*	RAD21

1 INTRODUCTION

The cohesin complex could be referred to as the "circle of life" as it underlies multiple key processes within a living cell and its functionality is imperative for survival. The name cohesin was coined due to the complex's role in the holding together of sister chromatids from S phase until G2 phase of the cell cycle, but it is now well established that, in addition to its cell cycle role, cohesin also works to organise 3-dimensional chromosome conformation and interactions with regulatory elements. Cohesin's influence in regulating gene expression opens doors to exploring this role not only in normal development, but in abnormal development and malignancies.

1.1 The cohesin complex

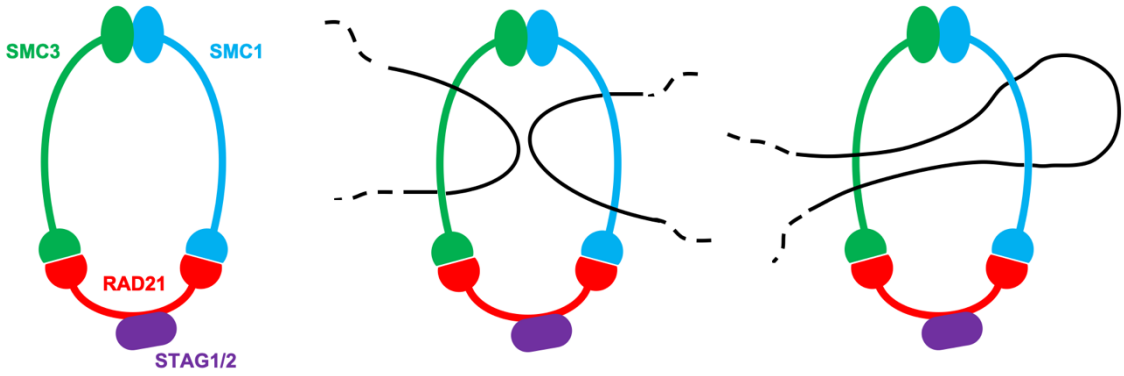


***Figure 1-1 - The cohesin complex**. Left-Schematic of the cohesin complex, with subunits annotated for orientation. Middle – Schematic of a cohesin complex holding two chromosomes in close proximity via the "handcuff" model. Right – Schematic of a cohesin complex holding two regions of a single chromosome together via the "loop extrusion" model.*

1.2 Cohesin structure and roles

The cohesin complex is a tripartite ring, containing two structural maintenance of chromosomes (SMC) subunits SMC1 and SMC3, and an α-kleisin subunit (Chavda et al., 2017; T. Gligoris & Löwe, 2016; Nasmyth & Haering, 2009). SMC1 and SMC3 are large proteins that join together at one end to form a hinge domain, leaving their ATPase heads to interact with the α-kleisin subunit (T. G. Gligoris, 2018; Haering et al., 2004; Nasmyth &

Haering, 2005). The α-kleisin interacts with stromal antigen subunits (STAG/SA/Scc3) (Viny et al., 2019; Ward et al., 2016)(Figure 1-1). The α-kleisin RAD21, also termed Scc1 or Mcd1 in yeast, is essential for cohesin's normal function, as there are no homologous proteins produced which can perform its role (Birkenbihl & Subramani, 1992; Sonoda et al., 2001; Tomonaga et al., 2000). The human RAD21 protein is 631 amino acids long and is encoded on chromosome 8 (Dorval et al., 2020). RAD21 not only links together SMC structures but is the only subunit of the cohesin complex which is noted to directly bind accessory proteins, recruiting them to carry out their function at cohesin-binding sites (Dorsett & Merkenschlager, 2013). The stromal antigen proteins, termed STAG1, STAG2 and STAG3 in humans, all have similar sequences, yet favour different roles within a cell. STAG3 is associated with meiotic cohesin, while STAG1 and STAG2 are ubiquitous (Casa et al., 2019; Prieto et al., 2001). Despite STAG3 being associated with meiosis 1, all STAG orthologues have the ability to bind to RAD21 in all cell types, so even a homozygous mutation for one of these genes is not lethal (Viny et al., 2019). The Human STAG2 protein is encoded on the X chromosome, therefore in males only a single hit to this gene is sufficient to knock out STAG2 function (Mullegama et al., 2019). Many cancers have a high STAG2 mutation rate, as though STAG1 can compensate for STAG2 function, the reduction in overall STAG protein may still affect normal cell division and gene expression (Romero-Pérez et al., 2019; Viny et al., 2019). Human STAG1 is encoded on chromosome 3 and STAG3 on chromosome 7 (Romero-Pérez et al., 2019).

Cohesin's loading onto chromosomes, maintenance and turnover are controlled by several other proteins (Nasmyth & Haering, 2009; Xiong & Gerton, 2010). It is proposed that cohesin is loaded and unloaded from chromosomes by a "dual gate" mechanism, in which cohesin snaps onto DNA via opening of the SMC hinge domains, and exits DNA via cleavage of the ring at the opposite end, at the a-kleisin interface (Nasmyth, 2011). There are a number of different models for how cohesin interacts with DNA, notably, the handcuff model and the sliding filament, or loop extrusion model (Figure 1-1). Locking chromosomes within a cohesin ring in a handcuff-like manoeuvre is the basis of the handcuff model (Eng et al., 2015; N. Zhang et al., 2008). The sliding filament, or loop extrusion model proposes instead that cohesin is loaded onto the DNA and slides along it, looping off regions of DNA

(Barrington et al., 2017; Davidson et al., 2019; Kim et al., 2019). Once loaded onto chromosomes, cohesin binds to DNA with variable degrees of permanency (Rhodes et al., 2017; H. Zhang et al., 2019), and has the ability to move along chromosome arms (Lengronne et al., 2004; Ocampo-Hafalla et al., 2016). During mitosis, cohesin is bound for long enough to fulfil its function in sister chromatid cohesion, due to acetylation by cohesin acetyl transferase (CoAT) proteins generating a stably bound form of cohesin along sister chromatids (Hou & Zou, 2005; Ivanov et al., 2018; Sun et al., 2019). Phosphorylation of cohesin assists in unloading cohesin as mitosis progresses, with centromeric cohesin the last of the stably bound cohesin to be removed (Hauf et al., 2005; Stormo & Fox, 2018). Stably-bound cohesin also works to hold homologous chromosomes together when a double-stranded break occurs, allowing repair machinery to use the intact chromosome as a template to repair the damage (Dorsett & Ström, 2012; Watrin & Peters, 2006). Cohesin complexes are enriched at damaged DNA sites (Hagstrom & Meyer, 2003), which supports the notion that the complex assists in holding chromosomes together while repair is undertaken. Cohesin's roles are not limited to assisting in cell proliferation and repair. In fact, cohesin is at the forefront of many modern studies because of its role in influencing gene expression (Dorsett & Ström, 2012; Richterova et al., 2017; Watrin et al., 2016). It is now well established that the cohesin complex can recruit activating or repressive regulators to distinct DNA sites to have their effect (Dorsett, 2019; Rahman et al., 2015). Newer studies have evidence for cohesin working alongside CTCF, a DNA binding protein that acts as a chromatin insulator, to regulate gene expression (Merkenschlager & Nora, 2016; Parelho et al., 2008; Pugacheva et al., 2020).

1.3 Topological organisation of the genome

Cohesin and CTCF work together to organise the 3-dimensional (3D) genome (Chavda et al., 2017; Chen et al., 2019; Kentepozidou et al., 2020). The CTCF protein contains 11 zinc finger proteins, which bind to side chains of exposed nucleotides in the major groove of consensus sequence DNA (Hansen et al., 2019). CTCF is the direct link between chromatin and cohesin at consensus sites, as the cohesin α-kleisin subunit directly binds to CTCF, which in turn is bound to consensus DNA (Pugacheva et al., 2020). CTCF knockdown results in a loss of cohesin recruitment to CTCF consensus sites, while knockdown of cohesin

resulted in a loss of CTCF's insulator function, with the resulting idea that both CTCF and cohesin are required to mediate insulated domains within local chromosome areas (Ali et al., 2016; Nora et al., 2017; Parelho et al., 2008). Cell-type-specific cohesin and CTCF binding supports the idea that these two complexes organise cell-type-specific gene expression patterns, depending on regulatory elements (activators or silencers) held within proximity of coding genes in cohesin and CTCF organised chromatin (Barrington et al., 2019; Mir et al., 2019).

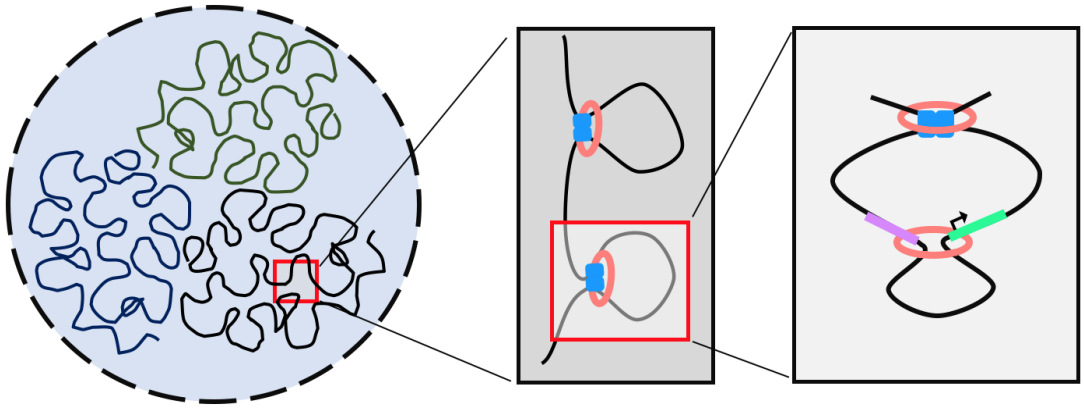


Figure 1-2- Schematic of nuclear 3D organisation*. Left - A nucleus, with distinct chromosome territories in different colours. Red square highlights a portion of a hypothetical active compartment. Middle - Close up of a region of the active compartment, showing 2 distinct TADs, each insulated by CTCF (blue) and cohesin (pink). Right - Close up of an insulated neighbourhood within one of the TADs, depicting CTCF and cohesin insulating a neighbourhood, while cohesin also works to organise local chromosome interaction through bringing a gene and enhancer together in a loop.*

Individual chromosomes segregate into "chromosome territories" throughout the nucleus, and A (active) and B (inactive) compartments bring together regions of mostly active and inactive transcription (Cremer et al., 2015; Fritz et al., 2016, 2019). Within A and B compartments, the genome is insulated into distinct chromatin domains, based upon whether single genes or groups of genes are actively or inactively transcribed. The largest of these regions are termed topologically associating domains (or TADs), in which cohesin and CTCF loop off a genomic region, resulting in a high frequency of interactions within that part of the genome (Barrington et al., 2019; Kentepozidou et al., 2020) (Figure 1-2-B). TADs are a large, tangible compartmentalisation of the chromosomes, in which interactions can occur, and just as importantly, interactions cannot occur between (Dixon et al., 2016;

Mir et al., 2019). These large TADs contain sub-TADs – which are simply smaller compartments of chromatin also insulated by cohesin and CTCF. Within sub-TADs are insulated neighbourhoods, smaller-again compartments insulted by cohesin and CTCF, where cohesin can entrap chromosomes to bring together regulators and genes at a local level (Figure 1-2-C) (Dowen et al., 2014; Hnisz et al., 2016; Sun et al., 2019). While TADs are conserved across related species, between different cell lineages the organisation of sub-TADs and insulated neighbourhoods differs, depending on the gene expression requirements of the cell type (Ali et al., 2016; Dowen et al., 2014). TADs across the genome are shown to be disrupted in mitosis, yet re-established straight after, consistent with the fact that cohesin localises to function in sister chromatid cohesion, then returns post-division to organise gene expression patterns (Hansen et al., 2017; H. Zhang et al., 2019).

1.4 Pathologies associated with cohesin dysregulation

As cohesin's many roles within a cell provide chromosomal integrity and the expression of essential genes, when cohesin levels are reduced, or lost, pathologies can occur. A total cohesin knockout is lethal, while a partial knockdown can be survivable, with a spectrum of phenotypes (J. Horsfield et al., 2012; Piché et al., 2019). The best known of these pathologies is Cornelia de Lange syndrome (CdLS), a broad-spectrum syndrome with multiple physical and cognitive abnormalities (Bergeron et al., 2019; Boyle et al., 2017; Mulder et al., 2017). CdLS patients are small in size and have a distinctive facial appearance, including arched eyebrows, long eyelashes, an upturned nose, thin upper lip, and an undersized jaw (Basel-Vanagaite et al., 2016; Yuan et al., 2019). Observed anomalies range from mild to severe, with the most severe having truncated limbs (Mehta et al., 2016). All patients within the CdLS spectrum have variable degrees of cognitive delay (Mikołajewska, 2019; Mulder et al., 2017). Nearly 70% of CdLS cases are caused by mutations in the NIPBL gene, one of the essential cohesin-loading proteins (Kaur et al., 2016; Masciadri et al., 2018). Mutations in SMC1 and SMC3 also give rise to CdLS phenotypes, and account for around 5% of cases. SMC-associated CdLS phenotypes are milder than those associated with NIPBL, though mental retardation is still present (Baquero-Montoya et al., 2014; Huisman et al., 2017). RAD21 mutations are causative of a cohesinopathy syndrome which partially phenocopies CdLS, though the observed

symptoms are even milder than patients with SMC mutations, presenting with mild cognitive and physical abnormalities (Deardorff et al., 2012; Dorval et al., 2020). Another rare human disorder, termed Roberts syndrome (RBS) is caused by mutations in ESCO2, a protein that acetylates cohesin, allowing adherence of sister chromosomes in mitosis (Banerji et al., 2017). The disorder is therefore due to premature separate of the chromosomes during mitosis (Tuuli & Odibo, 2018). Due to the numerous cohesin and cohesin factor mutations underlying these pathologies, the term "cohesinopathy" is used to describe these disorders. Most cases of cohesinopathy are heterozygous for mutations related to cohesin functionality, and the resulting loss of functional cohesin is predicted to provide enough cohesin to drive mitotic cell division yet be insufficient to drive expression of all of the essential developmental genes in a developing foetus, as there are so many, resulting in the abnormal phenotypes (Cucco & Musio, 2016; Tuuli & Odibo, 2018). Though it is theorised that the cohesinopathies all result from some degree of gene expression dysregulation, there are distinct differences in phenotypes observed when different components of the cohesin cycle are affected, indicating that the cohesin complex subunits have an affinity for regulating particular genes (Casa et al., 2019).

A loss of cohesin has also been attributed to myeloid malignancies, a group of haematopoietic diseases that are detrimental to survival. Myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML) are the two most common myeloid malignancies (Fisher et al., 2017; Fröhling et al., 2005; Shih et al., 2012). In MDS, hematopoietic stem cell function is compromised, resulting in dysplastic bone marrow (Gangat et al., 2016; Mhawech & Saleem, 2001), while in AML, myeloid progenitors do not differentiate into functional blood cells (De Kouchkovsky & Abdul-Hay, 2016; Saultz & Garzon, 2016). Compared to cancers derived from other tissues, myeloid cancers only require 5-15 mutations to drive malignancy (Farrar et al., 2016; Yoshida et al., 2013). While the mutational landscape of these malignancies is well characterised (Kandoth et al., 2013), the critical combination of somatic mutations leading to malignant progression is not yet established. Mutations in cohesin subunits are prominent in myeloid malignancies, occurring in 12.1% of acute myeloid leukaemias and 8.0% of myelodysplastic syndromes (Kon et al., 2013). Somatic cohesin mutations are implicated as a critical step in these

disease's development, as they tend to occur early within disease progression (Fisher et al., 2017; Hill et al., 2016; Losada, 2014). Patients with myeloid malignancy, yet a stable karyotype, are predicted to be a result of cohesin ineffectively controlling gene expression (Losada, 2014). Like with all malignancies, no myeloid dysplasia has been observed with a single mutation, and cohesin-subunit mutations are just one of a number of "hits" to the genome which drive malignancy for these patients.

1.5 Restoring runx1 expression to test Rad21 functionality

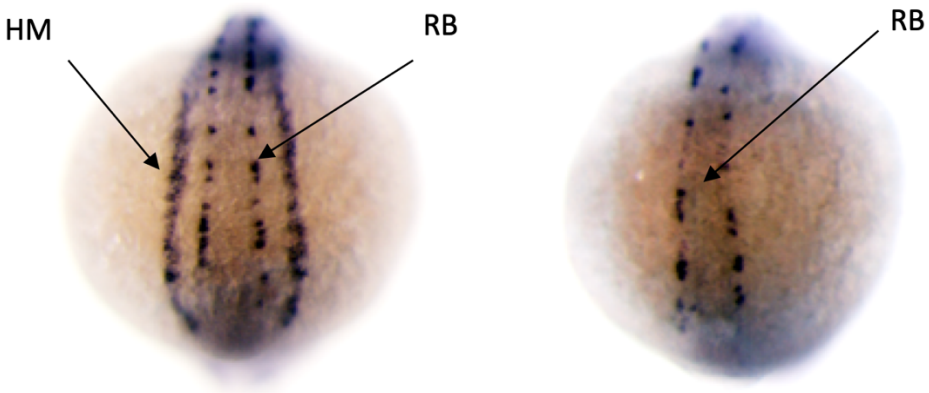


Figure 1-3 - Restoring runx1 expression to test rad21 functionality*. 13-hpf embryos imaged after runx1 in situ hybridisation. Haematopoietic mesoderm (HM) and Rohon Beard neurons (RB) labelled. Left - Embryo with genotypically-functional Rad21 and resulting runx1 expression in the RB and HM. Right - Embryo genotypically mutant for rad21. Exhibits loss of HM runx1 expression, while maintaining runx1 expression in the RB.*

The Horsfield lab has previously shown that *rad21*-homozygous mutant embryos exhibit differences in tailbud *runx1* expression relative to *rad21*-wildtype and *rad21*-heterozygote embryos at 13 hours post fertilisation (Figure 1-3). Embryos that produce functional Rad21 exhibit *runx1* expression in the haematopoietic mesoderm (HM), while Rad21-homozygous mutant embryos do not. (Horsfield et al., 2007). Microinjection of *rad21* mRNA into *rad21*-homozygous mutant embryos rescues HM *runx1* expression, producing embryos that present with wildtype-like *runx1* expression patterns at 13 hours post fertilisation (Horsfield et al., 2007). The functionality of human RAD21 mutations underlying cohesinopathies have been assayed through determining if analogous variant forms of *rad21*-mRNA could rescue *runx1* expression in the haematopoietic mesoderm of rad21-homozygous mutant embryos (Deardorff et al., 2012).

1.6 The canonical WNT pathway

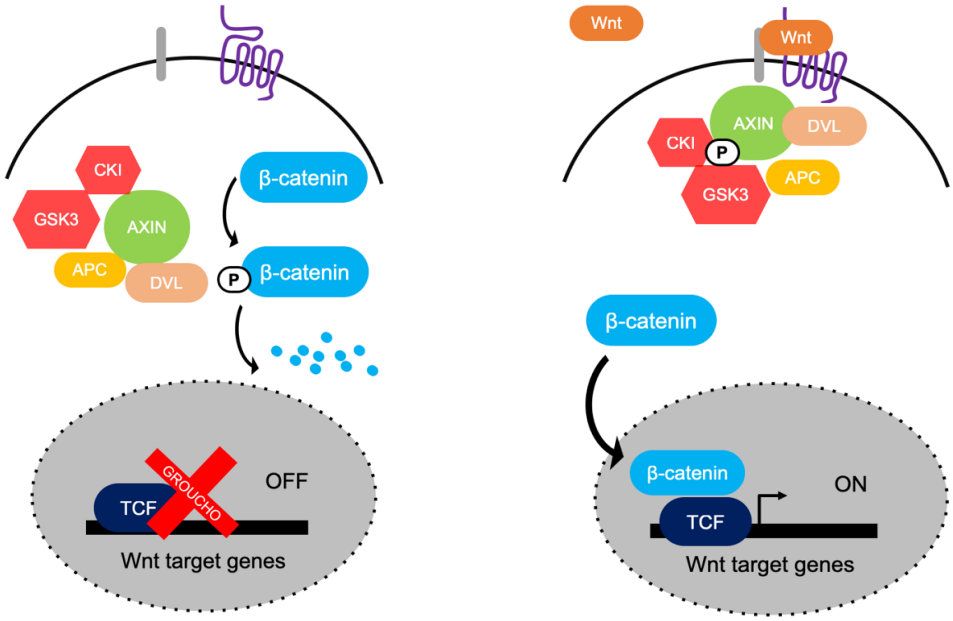


***Figure 1-4 - The canonical Wnt signalling pathway**. Image adapted from Wang et al 2004. Left - When Wnt receptor complexes are not bound by Wnt molecules, kinases CK1 and GSK3α/β phosphorylate β-catenin. Phosphorylated β-catenin is recognized by the ubiquitin ligase complex, which ubiquitinates it, targeting it for rapid destruction. In the nucleus, Groucho is bound to TCF, inhibiting the transcription of Wnt target genes. Right - Once bound by Wnt, receptor complexes activate the canonical signalling pathway. When Wnt binds, the destruction complex becomes disrupted. Axin and the destruction complex translocate to the receptor, and rather than target β -catenin for degradation, a series of phosphorylation events inhibit the GSK3α/β from doing so. Thus, β -catenin accumulates in the cytoplasm and translocates to the nucleus, where it can displace Groucho, and recruit other regulators to drive the expression of Wnt target genes*

The canonical Wnt pathway and Wnt target genes drive virtually every aspect of development, as well as maintaining adult tissues – making this field a hot spot for understanding both normal development and pathogenicity (Butler & Wallingford, 2017). The canonical pathway for Wnt signal transduction is outlined in Figure 1-4. When the pathway is activated by Wnt, β-catenin moves to the nucleus, where it displaces groucho from TCF sites, allowing the transcription of Wnt target genes (Wang & Wynshaw-Boris, 2004). These Wnt genes are responsible for axis patterning as well as cell specification, proliferation and migration (Darras et al., 2018; Xing et al., 2018). The canonical Wnt pathway is imperative in organising germ layers in the developing zebrafish tailbud. As the

posterior tailbud extends, bipotent stem cells make ectodermal or mesodermal decisions based upon their exposure to Wnt signalling (Martin & Kimelman, 2012).

1.7 Zebrafish as a model organism

Zebrafish are a popular animal model, for many reasons. Zebrafish are a genetically tractable vertebrate model on which cellular and molecular manipulation are possible. Zebrafish are small, easy to maintain, and can be kept in large numbers. Breeding between adult zebrafish can be performed once a week, and at each breeding, hundreds of fertilised eggs can be collected between a breeding pair. The eggs are fertilised outside of the adult fish, and can be collected from the water and taken away for further experiments and observation, which is simplified by the fact that the embryos are transparent.

1.7.1 Conservation of genes

Though the zebrafish genome is extensively rearranged, there is a great degree of conservation between the human and zebrafish genome, with nearly 70% of human genes having one or more zebrafish orthologues (Howe et al., 2013). Like humans, zebrafish have one copy of CTCF, Smc1, Smc3 and Rad21. Zebrafish do not possess a STAG3 orthologue, yet have additional copies of the two human stromal antigens, STAG1 and STAG2, termed Stag1a, Stag1b, Stag2a, Stag2b (Howe et al., 2013).

1.7.2 Zebrafish development

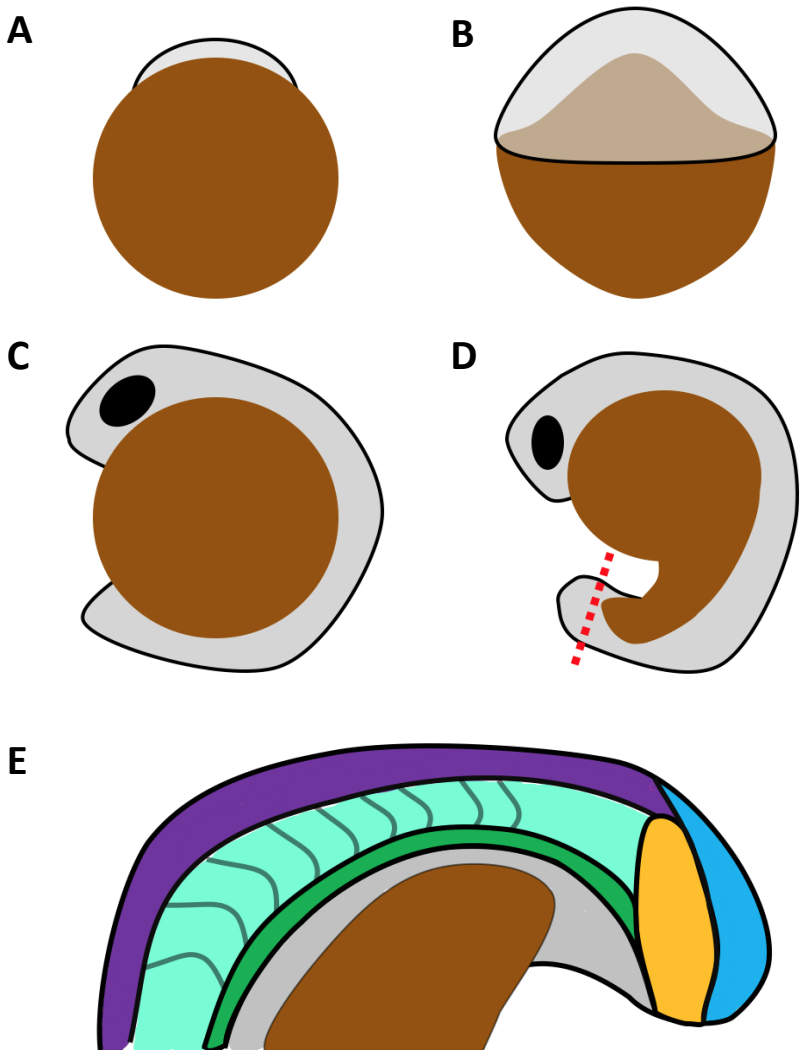


***Figure 1-5 - Key zebrafish developmental stages**. A -One-cell. B – 50% epiboly. C – 13 hours post fertilisation. D – 18 hours post fertilisation. Red dotted line marks site for tailbud excision (see method 3.2.14). E – Schematic of tailbud structures within the zebrafish tailbud. The tailbud is oriented A-P. Blue – Neuromesodermal progenitor population. Yellow – Mesodermal progenitor population. Purple – Neural ectoderm. Green – vascular endothelium. Teal – Presomitic mesoderm. Stripes in the presomitic mesoderm represent formed somites.*

Zebrafish develop from an embryo to a freely swimming fish over the course of a few days (Dooley & Zon, 2000). The zygote remains at the one-cell stage for almost 15 minutes, providing a short window of opportunity to inject constructs so that they will be present in all cells of the fish upon division (Figure 1-5-A). Somites and organs begin to form at ten

hours post fertilisation, and the tailbud begins to extend away from the yolk sac at around 13 hpf (Figure 1-5-C). By ~48hpf, the first fish are beginning to break free from their chorions and swim freely.

1.7.3 The zebrafish tailbud

The tailbud is an important structure that drives the posterior development of the organism. While primary gastrulation generates head structures, the posterior outgrowth of the tailbud, and a continual allocation to different cell fates from a bipotent cell pool at the very posterior of the tailbud, generates the posterior of the body axis (Smith & Kimelman, 2020). The neuromesodermal progenitor cells (NMP's) are bipotent stem cells at the very posterior of the tailbud (Figure 1-5-E) that escape early gastrulation patterning (Attardi et al., 2018; Martin, 2016). These cells remain in a bipotent state during somitogenesis, and specify into neural ectoderm and mesodermal progenitors, based upon their exposure to Wnt signalling (Dunty et al., 2008; Goto et al., 2017). Sox2, a neural fate driver, and Wnt proteins Wnt3a and Wnt8 are present in equal concentrations in the NMPs, with brachyury and Wnt engaged in a positive feedback loop to maintain levels of Wnt equal to that of Sox2, and thus maintain a bipotent state (Martin & Kimelman, 2012). Wnt signalling suppresses the neural fate and drives the mesodermal fate for differentiating NMPs (Gouti et al., 2014). When Wnt levels increase in an NMP the cell is directed to a mesodermal fate (yellow in figure 1-5-E), and when cells leave the NMP zone and are no longer exposed to Wnt signalling, they maintain Sox2 levels and commit to a neural fate (purple in figure 1-5-E) (Gouti et al., 2014, 2015; Koch et al., 2017). Cells continuously leave the NMP pool as the zebrafish elongates, adding structures to the posterior of the developing fish until the NMPs are exhausted and somitogenesis ends (Martin, 2016).

2 MASTERS OVERVIEW

This masters research comprises two projects. The first examines a familial RAD21 variant and seeks to answer whether this may be linked to observed familial myeloid dysplasia. The second examines the role of cohesin in regulating gene expression and cell fate decision making in the tailbud during zebrafish embryogenesis.

Both projects utilised the zebrafish strain *rad21*^{nz171}. This strain carries one truncated *rad21* allele, due to the G227* mutation present. As a total *rad21* knockout is lethal, these fish are held in heterozygous stocks in the Otago Zebrafish Facility, and in-crossed to produce wildtype, *rad21*^{nz171} heterozygous and *rad21*^{nz171} homozygous siblings.

2.1 Project one scope:

The aim of project one was to assay the functionality of the familial S449G RAD21 variant identified in the Australian SA Pathology myeloproliferative disorder cohort. The cohort of interest display myeloid dysplasia, and there is potential that this variant RAD21 protein could be linked to this prognosis. A Rad21 complementation assay similar to that performed in Deardorff et al 2012 will be performed in these analyses to assess if the familial variant mRNA produces functional protein and can rescue wildtype-like *runx1* expression in *rad21*^{nz171} homozygous embryos. Assigning function to the S449G-RAD21 variant is essential as it is a familial variant of unknown significance and may be underlying the observed myeloid dysplasia in the members of the cohort, with the possibility of causing dysplasia in subsequent generations of this family. Identification of S449G RAD21 functionality, or lack thereof, will bring these families one step closer to identifying what may be underlying their family history of malignancy. Alongside the familial variant, a novel RAD21 variant will be produced and assayed, giving an opportunity to examine another potentially pathogenic variant. Figure 2-1 illustrates the experimental overview

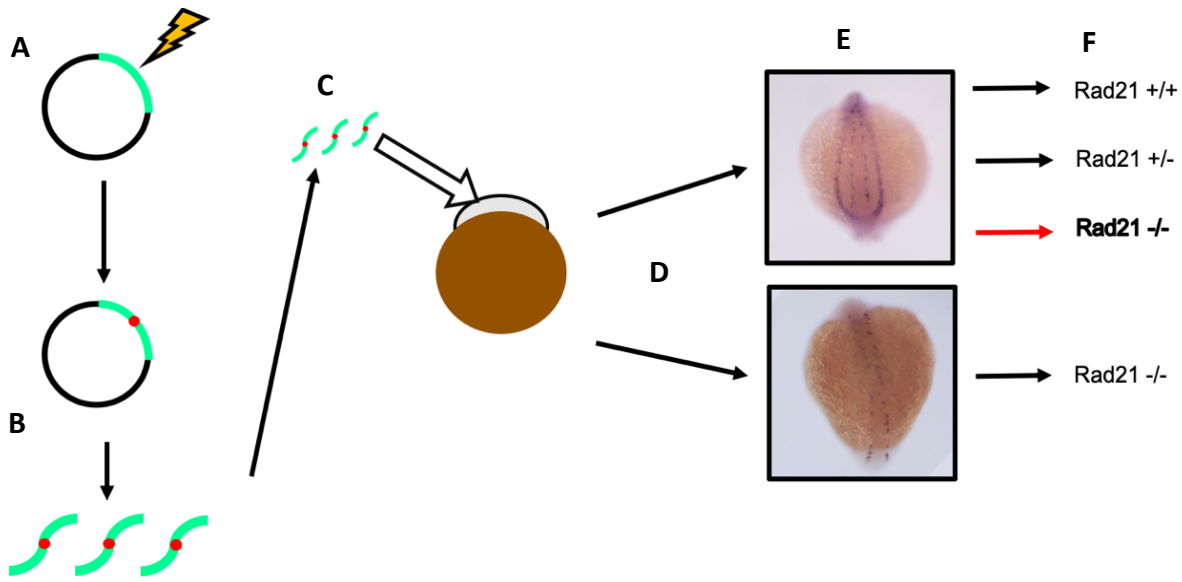


Figure 2-1 - Experimental overview for testing RAD21 functionality. *A – Create variant rad21 via site directed mutagenesis. B – Transcribe into mRNA. C- Inject variant mRNA into embryos from a rad21^{nz171} heterozygous in-cross. D – Perform runx1 in situ hybridisation. E – Observe phenotype. F – Confirm genotype.*

2.2 Project two scope:

The aim of project two was to identify the effect that cohesin has on regulating gene expression and cell fate commitment in the tailbud of the developing embryo, and how it may exert this effect. To identify all of the genes and related developmental pathways affected in a cohesin-deficiency, a high-throughput comparative RNA sequencing experiment was performed on the tailbuds of wildtype and cohesin deficient siblings from a *rad21^{nz171}* heterozygous in-cross. Gene expression in the elongating tailbud, which has not yet specified into distinct fates was examined with RNA-sequencing, as the posterior of the zebrafish develops from this point, making it a hub of transcriptional activity, and a lack of cohesin-regulated gene expression is predicted to cause dysregulation and abnormal cell fate decisions in this area. A number of the genes identified to be dysregulated in the RNA-sequencing analyses were examined by *in situ* hybridisation, to gain a more thorough picture of gene dysregulation in cohesin-deficiency. The spatial expression patterns of these genes were examined in siblings from a *rad21^{nz171}* heterozygous in-cross, and a *stag1b^{nz205}* heterozygous in-cross, to examine if gene expression phenocopying occurs when different components of cohesin are deficient. We predict that unconstrained

chromatin in cohesin mutants is open to ectopic signals, and thus regulators can have access to genes previously sequestered by chromatin organisation. In a bid to explore this idea, we utilised the canonical Wnt pathway as a model for all regulators and genes, and by agonising the pathway in cohesin-deficiency, we hoped to show that ectopic signals gain access to genes in unconstrained chromatin where they can be interpreted and produce aberrant gene expression.

3 MATERIALS AND METHODS

3.1 Chemicals and solutions

All chemicals were of analytical grade. All solutions are listed in Appendix 1. Solutions and dilutions were prepared using deionised Milli-Q water (Millipore Corporation) or UltraPureTM DNase/RNase-Free Distilled Water (Life Technologies) – here termed UltraPure water.

3.2 Methods

3.2.1 Zebrafish

Zebrafish (*D. rerio*) were housed and maintained in the Otago Zebrafish Facility. All experimental procedures were approved by the University of Otago Animal Ethics Committee and carried out in accordance with the Otago Zebrafish Facility Standard Operating Procedures.

3.2.1.1 Zebrafish strains used

The Rad21 mutant, *rad21*^{nz171}, Stag1b mutant, *stag1b*^{nz205}, and *rad21*^{nz171}Tg(7xTCFXla.Siam:nlsmCherry)ia5.

3.2.1.2 Maintenance and feeding

Each tank was kept at 24°C - 30°C, pH 6.5 - 8.5 and conductivity of 200 µS - 1000 µS. System water was purified using a ZebTEC system. Fish numbers were kept at or below 20 adults per 3.5 L tank. Fish were fed three times a day, consisting of two dry feedings and one live feed.

3.2.1.3 Breeding and collecting eggs

Adult zebrafish were set up in specialised breeding tanks, the afternoon prior to embryo collection. These tanks contain a false bottom and a removable barrier to separate male and females. The OZF has automated lighting, with lights turning on at 8 am. Removal of

the barrier once the lights come on in the morning allows the zebrafish to mate, as the innate breeding cue for zebrafish is light. To prevent the parents consuming any eggs, the false bottom of the tank allows fertilised eggs to slip through, where they can then be collected. Embryos are collected through a fine sieve, and transferred to a petri dish with E3 embryo medium supplemented with methylene blue, which acts as an anti-fungal agent.

3.2.1.4 Growing embryos

Embryos were grown in a petri dish in E3 medium supplemented with methylene blue. The rate of embryo development can be manipulated by incubating at either 28.5°C for normal development or slowing down development by incubating at 22.5°C. Embryos hatch naturally at 2-3 days post fertilisation, however the chorion can be removed earlier using #55 Dumont Forceps. When the chorions were discarded the media was changed to E3 without methylene blue.

3.2.1.5 Embryo fixation and storage

Embryos were collected at the desired developmental stages 13 hpf and 18 hpf (figure 1-5-C/D), and the chorions removed manually with forceps. Dechorionated embryos were then fixed in 4 % paraformaldehyde (PFA) in 1x PBS at 4 °C overnight. Fixed embryos were then dehydrated in 100 % methanol, and stored at −20 °C indefinitely.

3.2.1.6 Microinjection of zebrafish embryos

Microinjections of mRNA and morpholino were performed using a Narishige micromanipulator (UM-3C) and a MPPI2 Milli-Pulse Pressure Injector under a Leica MZ12 dissecting microscope. To microinject variant rad21 RNA, 100 pg of each mRNA was injected into the one-cell (figure 1-5-A) of embryos from a *rad21*^{nz171}heterozygote in-cross. 0.5 pmol of Rad21 morpholino was injected into the yolk sac of Tg(7xTCFXla.Siam:nlsmCherry)ia5 embryos at the 1-cell stage. Morpholino oligonucleotides (MOs) were obtained from GeneTools LLC and diluted in water. Rad21 MO sequence: 5'-AGGACGAAGTGGGCGTAAAACATTG- 3'.

3.2.2 Riboprobe synthesis

Maps of all plasmids are shown in Appendix 2. Bulk preps of plasmids containing genes of interest were created, and linearised with the appropriate restriction enzymes at 37°C for 2 h. Digestion was verified by electrophoresis on a 1 % Agarose gel. Linearised plasmids were purified via phenol/chloroform protein extraction. To create digoxigenin (DIG) labelled riboprobes, the Roche DIG RNA Labelling Kit was used. 1 µg of the purified plasmid DNA was used in a transcription reaction mix with either Sp6 or T7 RNA polymerase enzyme (Roche Diagnostics), 10x DIG labelled NTP (Roche Diagnostics) and 10x transcription buffer (Roche Diagnostics). The eluted riboprobes were quantified using the Nanophotometer and run on a 1 % agarose gel to check for quality. All probes were stored at −80 °C, and diluted 1:100 to give an appropriate probe concentration for in situ hybridisation readout.

3.2.3 Site directed mutagenesis of zebrafish rad21 constructs

The QuikChange II Site-Directed Mutagenesis Kit was used to mutagenise the zebrafish *rad21* gene within PCS2+ *rad21* (Appendix 2) after designing primers flanking appropriate regions. The kit was followed for generating point mutations, with 12 rounds of PCR amplification sufficient to incorporate these into the gene. Following amplification, 1 µL of Dpn I restriction enzyme was added to the amplification reactions, consistent with manufacturer's instructions, but cell transformation was adapted from manufacturer instructions (see method 3.2.4).

Sanger sequencing was performed by Genetic Analysis Otago (http://gas.otago.ac.nz/) to confirm that site directed mutagenesis of zebrafish Rad21 was successful. Primers are listed in Appendix 3.

3.2.4 Adapted cell transformation with rad21 constructs

5 µL of the SDM amplification reactions were added to 500 µL aliquots of XL1-Blue super-competent cells thawed on ice. The reaction was heat pulsed at 42°C for 45 seconds, then placed back onto ice for 2 minutes. 500 µL of SOC broth was added to the tube, then the tube was shaken at 37°C for 1 hour. Cells were pelleted by centrifugation at 3000 rpm for

10 minutes. All but 100 µL of supernatant was taken off, the pellet resuspended and cells plated onto an LB agar + Amp plate. Plates were stored at 37°C for 16 hours. Successfully grown colonies were then cultured in 1 mL LB + Amp overnight, and made into a glycerol stock by combining 750 µL of overnight culture with 250 µL glycerol, and storing at -80°C.

3.2.5 Plasmid extraction

PCS2+ variant-*rad21* plasmids were extracted from transformed cells with the Nucleobond Xtra Mini Kit for sanger sequencing identification of successful mutagenesis.

Bulk plasmid preps were performed with the Nucleobond Xtra Midi Kit for both plasmids containing variant *rad21* constructs – to create mRNA from, and plasmids containing genes of interest - for riboprobe formation (Appendix 2). Glycerol stocks for all plasmids were streaked onto a LB agar + Amp plate, and incubated at 37°C overnight. A single colony was then grown at 37°C in a shaking incubator in 5 mL LB + Amp the following night. A bulk prep was then grown the following night, by inoculating 100 mL of LB + Amp with 1 mL of the overnight culture. This bulk prep was processed with the Nucleobond Xtra Midi Kit, and plasmid concentrations measured on the NP80 Nanophotometer

3.2.6 RNA Synthesis

Variant and control *rad21* mRNAs were synthesised with an SP6 mMessage mMachine Transcription Kit according to the manufacturer's instructions. PCS2+ plasmids containing these *rad21* constructs (Appendix 2) were linearised with *Not I* restriction enzyme and purified with lithium chloride precipitation before mMessage transcription. All mRNAs were resuspended in ultrapure water. Successful mRNA production was visualised by running on a 4% agarose gel at 100 V for 20 minutes, and concentration confirmed with nanophotometer.

3.2.7 DNA extraction

DNA was extracted from adult fin clips and whole embryos by addition of the material to a PCR tube containing DNA extraction mastermix (16 µL water and 2 µL of 10x PCR buffer – MgCl (Applied Biosystems) per 1x reaction). 2 µL of 10 µg proteinase K was added to each

tube, and tubes were incubated at 50 minutes at 50°C, 10 minutes at 98°C, and then cooled to 4°C, where stored until required.

3.2.8 TaqMan genotyping

Following DNA extraction, the TaqMan *rad21*^{nz171} mutation assay was performed by adding 1 µL of DNA extraction mix to 19 µL of TaqMan mastermix (10 µL 2x KAPA hotstart polymerase, 8.5 µL H_20 and 0.5 µL 40x SNP mix (which contains two fluoro-probes, that bind to the different alleles respectively) per 1x reaction). The automated programme "Taqman_meier_kapa" was run on the LightCycler480 software to identify between mutant and wildtype alleles at the *rad21*^{nz171} locus. The mutant *rad21*^{nz171} allele binds the FAM probe (465-510nm) and the wildtype allele binds VIC (533-580nm), making a distinction between the genotypes clear.

3.2.9 Reporter line genotyping

Following DNA extraction, 1 µL of DNA extraction mix was added to 9 µL of PCR mastermix (1 µL of 10x PCR buffer (-MgCl), 0.3 µL of $MgCl_2$, 50 mM, 0.2 µL of dNTP mix, 10 mM, 7.06 µL of nuclease-free water, 0.04 µL of Platinum Taq DNA Polymerase, and 0.2 µL of each mCherry forward and reverse primer per 1x reaction) and run on the Thermocycler at 94 °C for 3 minutes, thirty cycles of : 94°C for 30 seconds, 57°C for 30 seconds and 72°C for 30 seconds, 72°C for 7 minutes, before cooling at 4°C indefinitely. The PCR product was then run on a 1% agarose gel for 120 V for 1 hour, and imaged using a GelDoc (Biorad). If a product was visualised, the fish carries the reporter gene.

3.2.10 HRMA genotyping

Following DNA extraction, 1.2 µL of DNA extraction mix was added to 9 µL of mastermix (2.55 µL water, 1.2 µL 25 mM $MgCl_2$, 5 µL 2x HRMA kit master mix and 0.125 µL of each forward and reverse primer per 1x reaction) and run on the LightCycler at 95 °C for 5 minutes, 40 cycles of: 95°C for 10 seconds, 60°C for 15 seconds, and 72°C for 20 seconds, before cooling to 37°C. Genotypes were interpreted from the melt curves, as the heterozygotes separate from the homozygotes distinctly, and homozygous mutants can be identified in comparison to homozygous wildtype controls. Primers listed in Appendix 3.

3.2.11 Whole-mount *in situ* hybridisation and preparation for imaging

For *in situ* hybridisation of embryos, both overnight hybridisation and wash steps were performed at 65°C. Washes were performed as follows: 25% 2×SSCT/75% hybe-buffer, 50% 2×SSCT/50% hybe-buffer, 75% 2×SSCT/25% hybe-buffer, 100% 2×SSCT (15 minutes), 0.2×SSCT (2x30 minutes), PBST (3x5 minutes). Embryos were blocked in 2% blocking reagent for 2-3 hours. Riboprobes were detected with anti-DIG antibodies coupled to alkaline phosphatase (AP). Excess antibody was removed by 8x15 minute PBST washes. Bound antibody was visualised using the AP substrate NBT/BCIP. Once the stain developed, embryos were washed 3x5 minutes in PBST to stop further stain development, then fixed in PFA overnight. PFA cleared with 2x5 minute PBST washes, then the prepared embryos stored in 80% glycerol at 4°C until ready for imaging.

3.2.12 Imaging *in situ* embryos

Embryo images were captured on a Leica M205FA microscope using a Leica DFC490 camera linked to Leica Application Suite V4 software. A dimpled slide was used to manipulate embryos in glycerol for imaging. As embryos could not be sorted into their respective genotype by morphology alone, they were genotyped post-imaging.

3.2.13 Stimulating embryos with WNT agonist

Rad21^{nz171} Tg(7xTCFXla.Siam:nlsmCherry)ia5 fish were in-crossed and the collected embryos grown until 50% epiboly, at which the chorion was torn, and embryos added to the well of a 24-well plate and immersed in 600 µL of E3 supplemented with 2.5uM of BIO. The embryos were then grown until 18 hpf, at which point their mCherry fluorescence was examined with the Nikon C2 confocal microscope, and associated NIS imaging software.

3.2.14 Tailbud excision

rad21^{nz171} heterozygote fish were in-crossed, and the embryos grown until 18 hpf. Each embryo was manually dechorionated with Dumont 55 forceps, and using the forceps, the tailbud was dissected away from the rest of the embryo. The dissected tailbud was dissolved in 3 µL of RLT + BMe (Qiagen RNeasy) in a PCR tube, and stored at -80 °C, while

the top half of the embryo was placed in 18 µL of DNA extraction mix (see method 3.2.7) in a separate PCR tube with a matching label. DNA was extracted from the top half of the embryo, and TaqMan genotyped, providing the genotype of the stored tailbud, as genotypes cannot be inferred from morphology at 18 hpf. Sibling wildtype and *rad21*^{nz171} homozygote embryos were used for comparison, and 4x pools of 80 tailbuds were collected for each genotype. *Rad21*^{nz171} heterozygote tailbuds that were inadvertently collected were not used for analysis.

3.2.15 RNA extraction

RNA was extracted from pooled tailbud samples using the Qiagen RNeasy Micro kit. The manufacturer's instructions were followed, with some optimisation of the process. All buffers were warmed to 23°C before use, and buffer RDD was warmed to 30°C before use. During the RPE wash step, tubes were rolled horizontally back and forth for 30 seconds to ensure guanine salts (present in RLT) would not contaminate the final elution. Samples were analysed by Qubit, Bioanalyser, and nanophotometer to ensure no there was no RNA degradation and that sufficient quantities were present for sequencing.

3.2.16 RNA-Sequencing

RNA samples comprising four pools per condition from *rad21*^{nz171} homozygote and wildtype tailbuds (n=80 per pool) were sequenced to compare gene expression profiles. TruSeq stranded mRNA libraries were prepared by Genomics Otago, and paired end 2x50 base pair sequencing was performed on an Illumina HiSeq 2500 V2. Between 40 and 55 million reads were generated for each replicate.

3.2.17 RNA-sequencing analysis

Gregory Giminez assisted with analyses. Using FastQ-Mcf tools, the reads were adaptor and quality trimmed and sequences with a Phred score lower than Q20 were removed. The cleaned reads were aligned against zebrafish genome version GRCZ11 using HISAT2 version 2.0.5. The reference genome was indexed before alignment. Read counts were retrieved first by exon and then summarised by gene using FeatureCount, only reads with a mapping quality of 10 were taken into account to avoid spurious alignment. Euclidean distance

matrices and PCA analysis confirmed clustering of distinct conditions. A normalised Bland Altman MA plot identified dysregulated genes, and reactome and gene ontology pathways analyses identified dysregulated pathways between the conditions. Identities for the top 50 most dysregulated pathways were manually plotted to provide visualisation of dysregulation in *rad21*^{nz171} homozygote samples.

3.2.18 Protein prediction

PyMOL and STRUM software was used for protein alignments and an annotate script written by David Markie (Department of Pathology, University of Otago) which utilised BCFtools (v1.9) adding metrics Polyphen-2, MPC, REVEL and PrimateAI, was used to predict pathogenicity of *rad21* variants.

3.2.19 Other software

Biorad GelDoc XR+ with Quantity One imaging software used for all gel images. NP80 Nanophotometer was used to assess RNA and DNA concentrations and purities throughout.

4 ASSAYING FUNCTIONALITY OF RAD21 VARIANTS

Injections of variant *rad21* mRNA into *rad21*^{nz171} homozygote embryos and subsequent *runx1 in situ* hybridisation at 13 hours post fertilisation allows identification of the functionality of the variant, as a functional *rad21* mRNA will rescue *runx1* expression in the haematopoietic mesoderm, while non-functional *rad21* mRNA will not (Deardorff et al., 2012). The functionality of familial S449G-RAD21 was assayed alongside a novel variant that was designed for these analyses, I152F-RAD21, which has not been identified in any cohorts to date. The novel variant I152F was produced at the opposite end of the *RAD21* gene to S449G and introduced a large phenol group to the protein, in a bid to create protein folding changes or steric hindrance at the binding interface with SMC3.

4.1 Production of variant *rad21* mRNA

Variant status	**Appropriate nucleotide change**	**Resulting codon change**	**Amino acid change in human RAD21**
Novel	A742T	**A**TC > **T**TC (I>F)	**I152F**
Familial	A1346G	**A**GU > **G**GU (S>G)	**S449G**

Table 4-1 - Creating amino acid changes in RAD21 via a single nucleotide change*. The single nucleotide variation, along with the codon and amino acid changes that result are outlined for both the novel and familial variants.*

Table 4-1 shows both the familial and novel variants, and which amino acids change because of these single nucleotide variations. To maintain consistency throughout, the variants will be referred to by their amino acid changes in the human *RAD21* gene, I152F and S449G.

Figure 4-1 - Creating human RAD21 mutations in the zebrafish rad21 gene*. Top: Alignment of human RAD21 and zebrafish rad21, with I152F and S449G annotated in orange. Middle: Variant-flanking primers designed to incorporate the point mutations, shown as green annotations spanning the orange sites of variation. Bottom: Alignment of rad21 and sequencing files for S449G-rad21 and I152F-rad21 after successful site-directed mutagenesis.*

Human *RAD21* and zebrafish *rad21* were aligned in Geneious and the I152F and S449G variants of interest annotated at the analogous sites in *rad21 (*Figure 4-1 – Top). Primers were designed for the PCS2+-*rad21* plasmid (Appendix 2) flanking the sites which were to have the appropriate point mutations inserted, and site directed mutagenesis (SDM) performed.

4.1.1 Optimisation of SDM for the S449G-*rad21* variant

Sanger sequencing confirmed that the SDM process was successful in producing the I152F variant in *rad21* immediately. However, SDM had to be optimised for the S449G variant, as the kit-directed twelve rounds of PCR amplification did not incorporate the S449G mutation. Sixteen rounds of amplification were performed with a gradient of annealing temperatures, and 57 °C was found to be optimal for the PCR reaction. Subsequent Sanger sequencing confirmed the production of the S449G-*rad21* variant.

Successful production of variant plasmids allowed for control and variant *rad21* mRNA transcription. The mRNA integrity was confirmed as suitable for the *runx*1 in situ bioassay (Appendix 4).

4.2 Assaying phenotype and confirming genotype

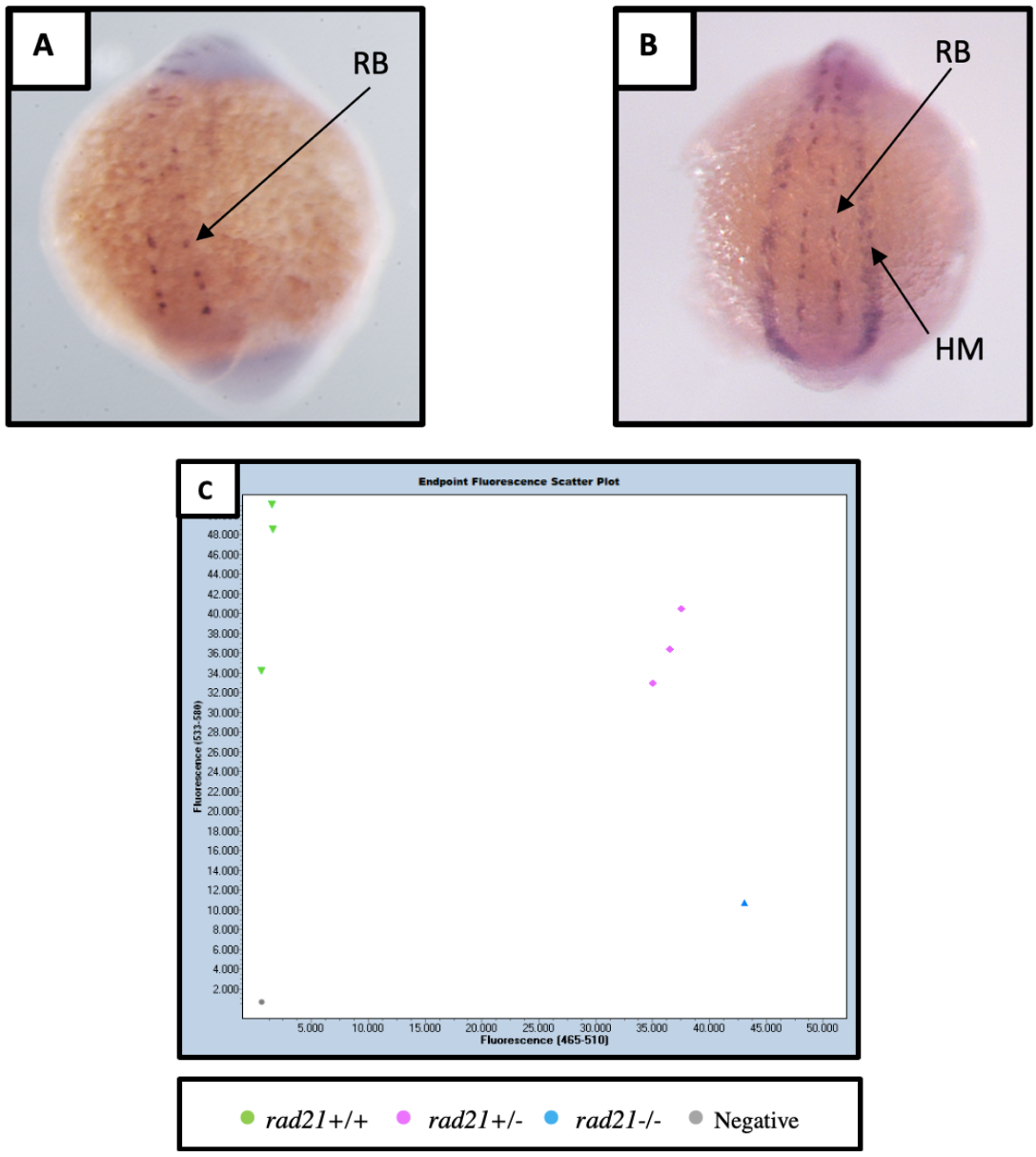


Figure 4-2 - Assaying rad21 phenotype and genotype*. A – Representative image of a rad21^{nz171} homozygote embryo following runx1 in situ hybridisation at 13 hpf. B – Representative image of a wildtype-like embryo following runx1 in situ hybridisation at 13 hpf. Runx-1 is expressed in Rohon-Beard neuronal progenitors (RB) and haematopoietic mesoderm (HM). C - TaqMan analyses chart demonstrating clear distinction between wildtype (rad21+/+), rad21^{nz171} heterozygote (rad21+/-), and rad21^{nz171} homozygote (rad21-/-) samples at the rad21^{nz171} variant site in rad21 (as labelled). Negative control also shown. The mutant allele binds to the 465-510 fluorophore (x axis) while the wildtype allele binds to the 533-580 fluorophore (y axis).*

A 1:100 riboprobe dilution produced specific *runx1* staining with little background, which is imperative for effective phenotyping of embryos. *rad21^{nz171}* homozygote embryos express

runx1 in the Rohon Beard neurons (RB), but not in the haematopoietic mesoderm (HM) (Figure 4-2-A). Both wildtype and *rad21*^{nz171} heterozygote embryos express *runx1* in the RB and the HM, an expression pattern termed "wildtype-like" for these analyses (Figure 4-2-B). TaqMan genotyping distinguishes between wildtype and *rad21*^{nz171} heterozygote embryos, which cannot be distinguished phenotypically (Figure 4-2-C). Once *rad21* mRNA microinjections are performed, and functional Rad21 protein produced, *runx1* staining in the HM of *rad21*^{nz171} homozygote embryos is rescued, and they too have a wildtype-like phenotype. TaqMan genotyping is therefore imperative to distinguish if wildtype-like embryos are producing their own Rad21 protein, or gaining Rad21 functionality through microinjected *rad21* mRNA. Due to Mendel's law of independent assortment, only 25% of the embryos from the *rad21*^{nz171} heterozygous in-cross will be *rad21*^{nz171} homozygous, and therefore informative in the assay.

4.3 Assaying RAD21 functionality

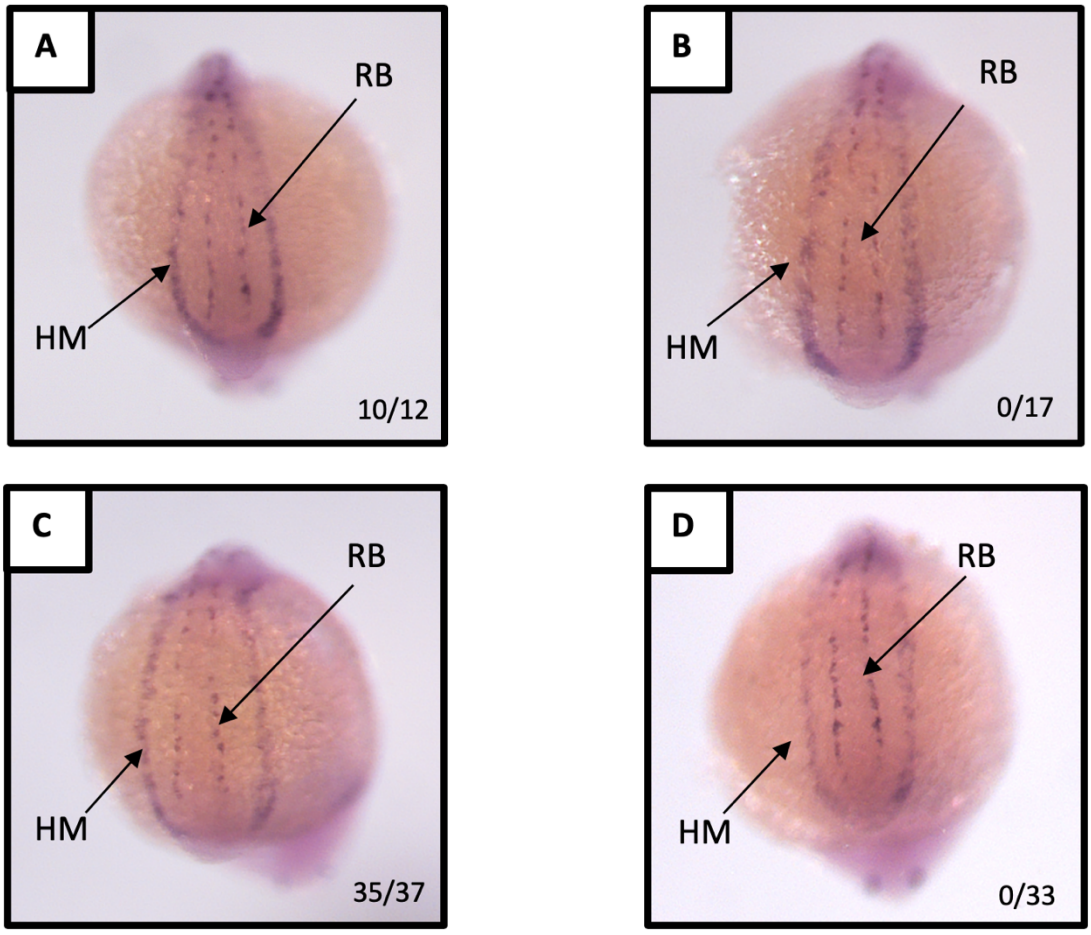


Figure 4-3 - Determining runx1 rescue in rad21^{nz171} homozygote embryos*. Representative images of wildtype-like rad21^{nz171} homozygote embryos following injection with rad21 mRNAs at the 1-cell stage, followed by runx1 in situ hybridisation at the 13 somite stage. A - rad21 injected embryos, B - rad21^{nz171} injected embryos. C- S449G-rad21 injected embryos, D- I152F-rad21 injected embryos. Numbers in the bottom right reflect the number of rad21^{nz171} homozygote embryos that produced wildtype-like runx1 expression due to functional rad21 mRNA injection. Rohon-Beard neuronal progenitors (RB) and haematopoietic mesoderm (HM) runx1 expression labelled.*

Rad21 and *rad21^{nz171}* mRNA were microinjected as positive and negative controls respectively, to establish the Rad21 complementation assay. *Rad21* injections restored HM *runx1* staining in 10/12 *rad21^{nz171}* homozygote embryos while *rad21^{nz171}* injections could not rescue *runx1* expression in any of the 17 *rad21^{nz171}* homozygote embryos injected, confirming that the bioassay was effective in determining Rad21 functionality . The two *rad21^{nz171}* homozygote embryos injected with wildtype *rad21* mRNA that did not exhibit *runx1* rescue are predicted to be due to microinjection error.

Variant *rad21* mRNA was assayed and rescue determined. S449G-*rad21* was effective in rescuing *runx1* HM expression in 35/37 *rad21*^{nz171} homozygote embryos (Figure 4-3-C), with the 2 *rad21*^{nz171} homozygote embryos not rescued predicted to be due to microinjection error. Successful HM *runx1* rescue shows that the S449G variant is functional, at least for rescuing *runx1* gene expression. I152F-*rad21* mRNA did not rescue *runx1* HM expression in the 33 *rad21*^{nz171} homozygote embryos injected, indicating that this novel variant is non-functional during development (Figure 4-3-D).

4.4 Predicting 3D protein structure of the variants

The novel I152F variant was designed to result in an isoleucine to phenylalanine substitution, which changes the properties of the amino acid from aliphatic to aromatic, as it adds a large phenol group to the structure. The familial S449G conversion does not drastically change the size of the amino acid, but changes the site from hydrophilic to hydrophobic, which could affect how the protein folds in 3D. If either of the variants amino acid changes result in the final protein folding differently, this could affect the ability of RAD21 to interact with other cohesin subunits, and compromise cohesin functionality. Currently, there is no complete RAD21 crystal structure. STRUM analyses generated a proposed 3D structure of the variant RAD21 proteins, to determine if the I152F or S449G variant sites lie within binding or active domains of the RAD21 protein. The two STRUM predictions have an identical 3D shape (Appendix 4), which indicates that the amino acid changes do not result in dramatic protein folding changes. The STRUM predictions indicated that I152F may be located within one of the SMC3 ATPase binding domains (Appendix 4). Steric hindrance at the ATPase binding site could disrupt cohesion complex formation, explaining the observed lack of rescue in the functional analyses.

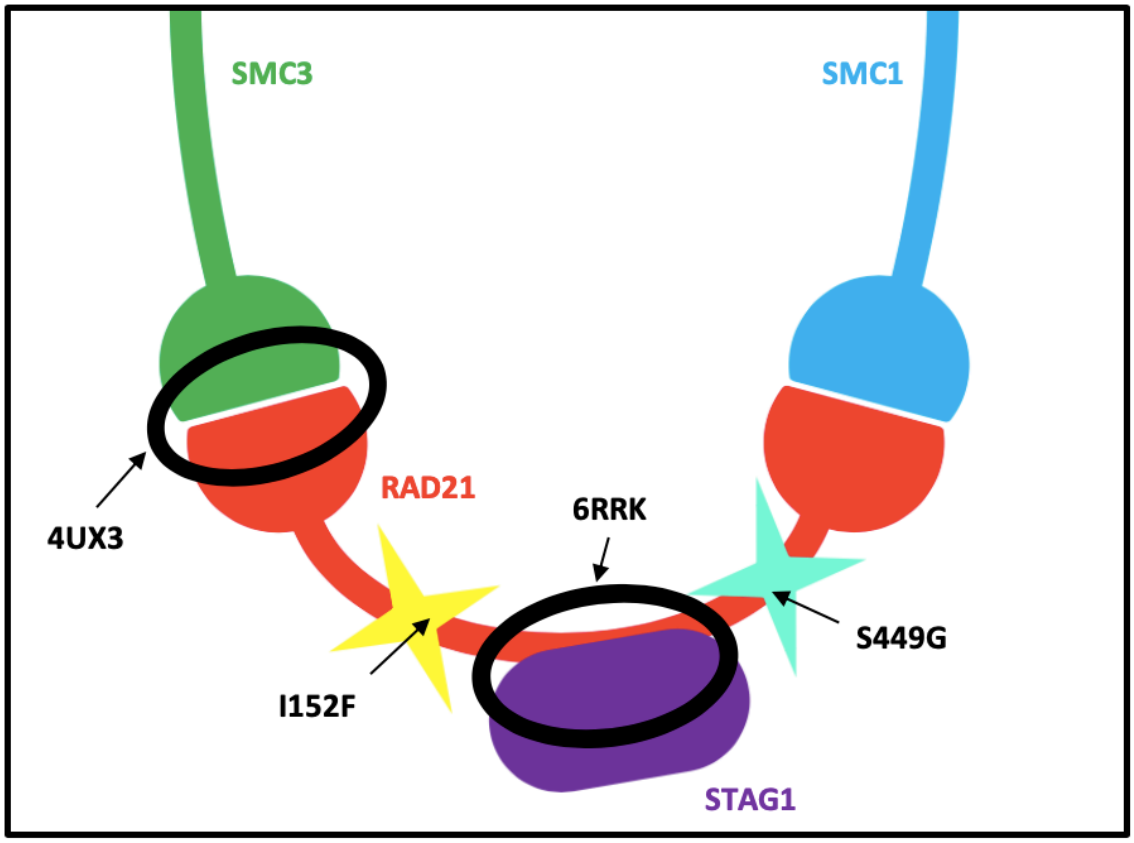


Figure 4-4 - Schematic of proposed variant sites within RAD21*. Schematic of Rad21 interactions with other cohesin subunits. Existing RAD21 crystal structures 4UX3 and 6RRK are circled in black. Predicted variant sites within Rad21 labelled with yellow and teal stars.*

There are several 3D structures in the Protein Data Base (PDB) that contain fragments of RAD21, where it is bound to other cohesin subunits. These structures were aligned to the variant RAD21 amino acid sequences in PyMol, to establish if the I152F variant interacts with the SMC ATPase head, as the STRUM prediction indicated. It was clear that it did not. Figure 4-4 depicts the predicted variant sites in relation to the known crystal structures. 4UX3 is an existing structure of the SMC3 ATPase head bound to RAD21, and 6RRK is the STAG1 subunit in complex with the middle portion of RAD21. The variants of interest are not contained within these known fragments of RAD21, but with mapping of their amino acid position it was determined where they may be located within the 3D protein. I152F is located towards the SMC3 binding domain, but between where SMC3 binds and where STAG binds. S449G is located towards the SMC1 binding domain, but between where SMC1 binds and where STAG binds. The variant positions are indicated by stars in figure 4-4.

Prediction programme	What it shows	S449G	I152F
PolyPhen-2	Scale from benign to possibly damaging	Possibly damaging	Benign
MPC	How "bad" a missense variant is a 0-5 scale (5 being very bad), scores >2 are enriched for disease causing variation	0.4 (Benign)	2.3 (Likely disease causing)
REVEL	This one is extremely conservative, with a score >0.6 you can be pretty sure it is doing something bad	0.1 (Benign)	0.08 (Benign)
Primate AI	Scores >0.8 are typically used as a cut-off for pathogenicity	0.46 (Benign)	0.85 (Pathogenic)

Table 4-2 - Pathogenicity predictions from a number of different programmes.

The variant amino acid sequences were analysed with a number of different pathogenicity prediction pipelines, and the outcomes are depicted in Table 4-2. Most of the pipelines predicted that S449G-RAD21 would be benign, yet there were conflicting results amongst the different programmes about the pathogenicity of I152F-RAD21. PolyPhen-2 and REVEL predicted I152F to be benign, while MPC and Primate AI predicted it to be pathogenic.

4.5 Conclusion

This experimental design was used due to its previous success in uncovering RAD21 functionality in a developmental context (Deardorff et al, 2012). It was important to discover the functionality of the familial S449G-RAD21 variant as it is of unknown significance, yet may play a critical role in the development of myeloid dysplasia, which is observed in this family. We have determined that S449G-RAD21 is functional, at least in early developmental stages, as analogous zebrafish S449G-*rad21* mRNA was sufficient to rescue a wildtype-like phenotype in *rad21*^{nz171} homozygote embryos. The novel variant I152F-RAD21 does not produce a functional cohesin subunit, as no haematopoietic mesoderm *runx1* expression was rescued in *rad21*^{nz171} homozygotes injected with I152F-

rad21 mRNA. Mapping existing fragments of RAD21 allowed identification of a generalised location of the variant sites within the 3D protein. Neither of the variants appear to be located in direct binding domains of RAD21 to other cohesin subunits, but this does not mean that they are not capable of causing steric hindrance in cohesin complex formation. Pathogenicity programmes and STRUM, a protein prediction programme, both highlighted that though algorithms are a great indicator of what may be happening biologically, they are not as accurate and as informative as functional assessment. Taken together, the evidence supports S449G-RAD21 as functional, and I152F-RAD21 as non-functional.

5 COHESINS ROLE IN REGULATING DEVELOPMENTAL GENE EXPRESSION

To gain a broad understanding of gene expression patterns and cell fate decisions in cohesin-deficient development, a tissue-specific RNA-sequencing experiment was performed, analysing genes expressed in the developing zebrafish tailbud. In this experiment, averaging effects across the whole embryo would not interfere with sequencing readout, as they may have in previous studies (Meier et al., 2018). The tailbud is a fascinating component of the zebrafish embryo, as it elongates away from the yolk sac as development progresses, contributing to multiple essential structures. We aimed to identify which genes and pathways cohesin controls in the developing zebrafish tailbud.

5.1 RNA-sequencing experimental overview

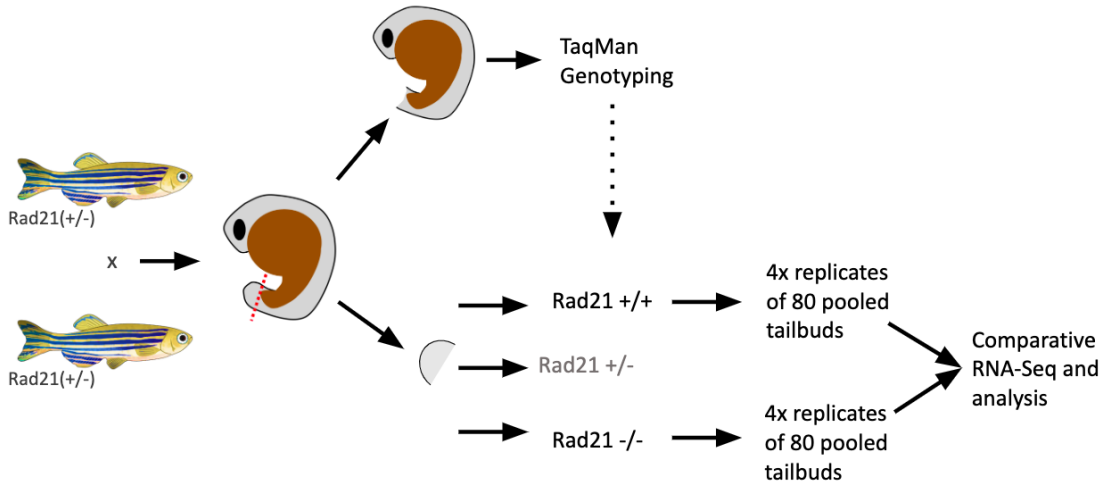


Figure 5-1 - Outline of the comparative RNA-sequencing experiment*. Genotypes denoted on diagram as follows: Wildtype – Rad21+/+, Rad21^{nz171} heterozygote – Rad21 +/-, Rad21^{nz171} homozygote – Rad21 -/-. Rad21^{nz171} heterozygote embryos in-crossed and the embryos grown until 18 hours post fertilisation. At 18 hours post fertilisation, the tailbud was dissected away from the embryo, and stored in a labelled tube. Genotypes cannot be inferred from morphology at this stage, therefore the top half of the embryo was TaqMan assayed to identify the genotype of the stored tailbuds, which were then pooled in groups of 80, 4x replicates. RNA-sequencing and comparative analysis performed between wildtype and Rad21^{nz171}homozygote siblings.*

Sample	Concentration (ng/µl)	RIN number
Rad21^{nz171} homozygote 1 (M1)	113.7	9.1
Rad21^{nz171} homozygote 1 (M2)	79.5	9.1
Rad21^{nz171} homozygote 1 (M3)	97.8	9.0
Rad21^{nz171} homozygote 1 (M4)	81	8.8
Wildtype (WT1)	80.1	8.8
Wildtype (WT2)	79.5	9.7
Wildtype (WT3)	75.6	9.1
Wildtype (WT4)	99.9	6.5

Table 5-1 - RNA quality measurements*. Samples are listed with respective concentrations and quality measurements.*

The process of collecting wildtype and cohesin deficient tissue for examination is outlined in Figure 5-1. Any *rad21*^{nz171} heterozygote tissue that was inadvertently collected during the process was not sequenced in these analyses. Preliminary mRNA quality data is outlined in Table 5-1. Sufficient quantities of RNA were extracted, and this RNA was not degraded, as determined by the high RIN numbers, making it suitable for RNA-sequencing.

5.2 Confirming read depth

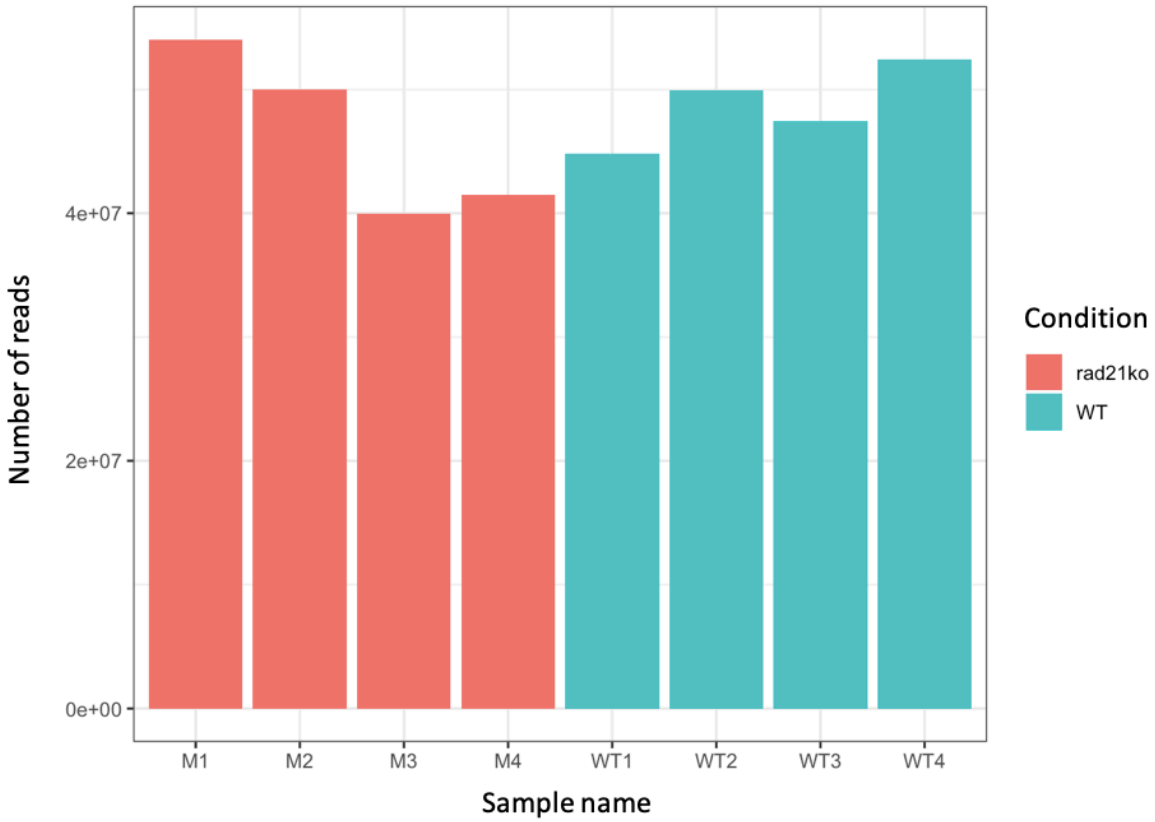


Figure 5-2 - Bar graph depicting the number of reads per RNA-sequencing replicate

All replicates had consistent read depth, which highlighted that the experimental work preceding the analyses was clean and that despite the cohesin-deficiency, mutant embryos still efficiently express genes . With at least 40 million reads per replicate, there was ample gene expression information for further analyses.

5.3 *Rad21* expression and verification of mutation

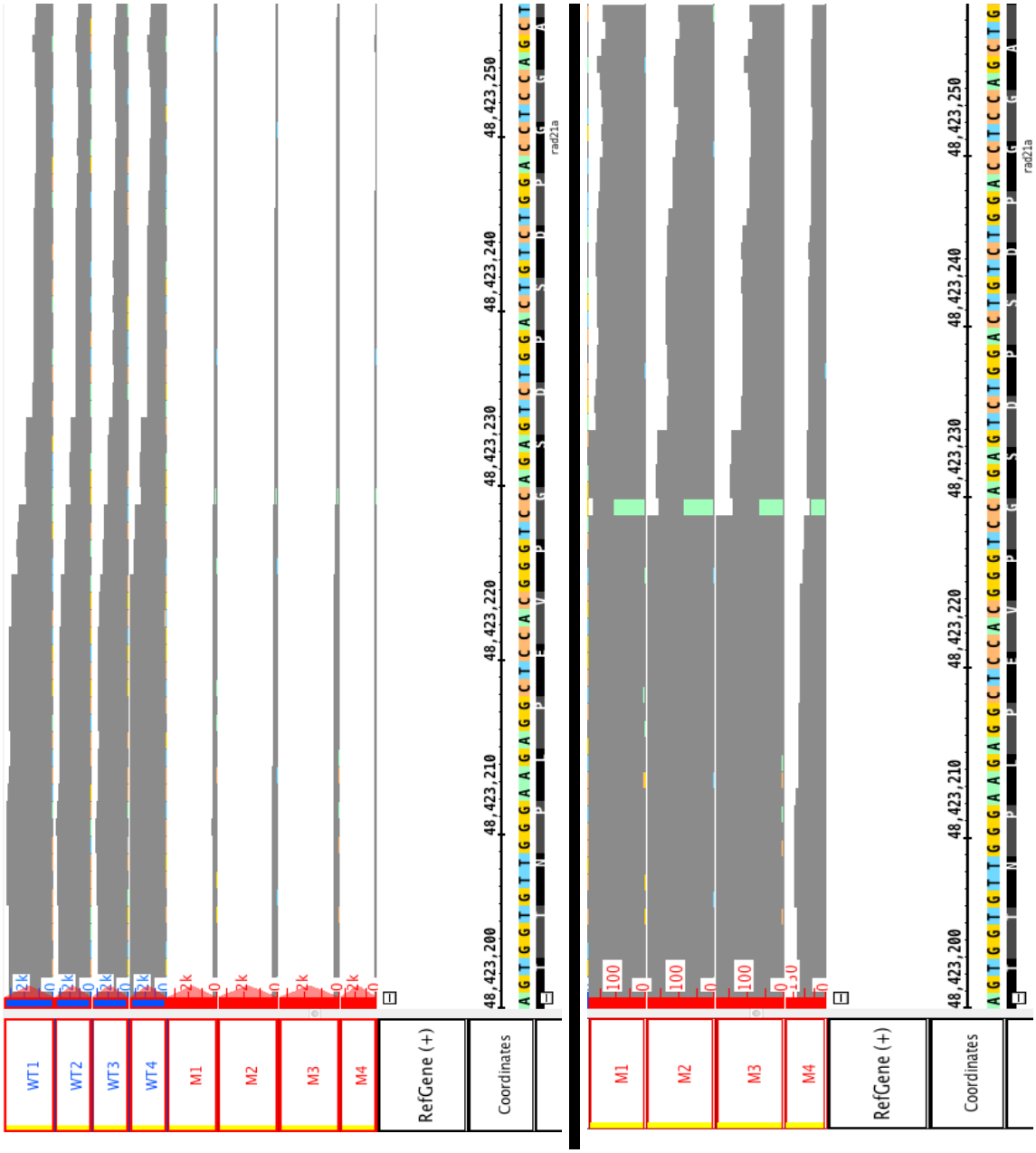


Figure 5-3 - Confirming loss of rad21 expression in rad21^{nz171} homozygotes*. Both panels depict the 60 nucleotides surrounding the G>T base pair change which produces the null rad21 allele. Top panel – Expression of rad21 in wildtype (blue) and rad21^{nz171} homozygote (red)samples. Grey histogram depicts read depth. Bottom panel – A close up of the read depth surrounding the G>T base pair change which produces the null rad21 allele, in rad21^{nz171} homozygote samples. Grey depicts the functional rad21 allele read depth and green depicts the null rad21 allele read depth.*

It is clear that *rad21* expression is greatly reduced in *rad21*^{nz171} homozygote samples compared to wildtype. (Figure 5-3 – Top panel). There are at least 3000 copies of *rad21* mRNA in each of the wildtype replicates, yet less than 300 in each of the *rad21*^{nz171} homozygote replicates. Upon closer examination of the small amount of *rad21* transcript present in the *rad21*^{nz171} homozygote samples, it became apparent that there is residual maternally deposited *rad21* present, as the *rad21*^{nz171} homozygote embryos cannot produce the functional allele, only the null allele, of which small amounts of transcription were also observed (Figure 5-3 – Bottom panel). The presence of maternally-deposited *rad21* mRNA reflects previous findings in which *rad21* mRNA was detected by RT-PCR in *rad21*^{nz171} homozygotes early in embryogenesis and immunoblotting showed Rad21 presence in *rad21*^{nz171} homozygotes at 18 hpf (Horsfield et al., 2007). The low levels of *rad21*^{nz171} mRNA present in the homozygotes is likely due to nonsense mediated mRNA decay, as the null allele produces a stop codon in the gene. The exact amount of functional *rad21* transcript present in the *rad21*^{nz171} homozygote samples was between 0-74 copies. As these were pools of 80 tailbuds it is evident that there is almost negligible maternal *rad21* left at 18 hpf and it is therefore likely inefficacious on gene expression from this time point in development.

5.4 Confirming distinction between the samples

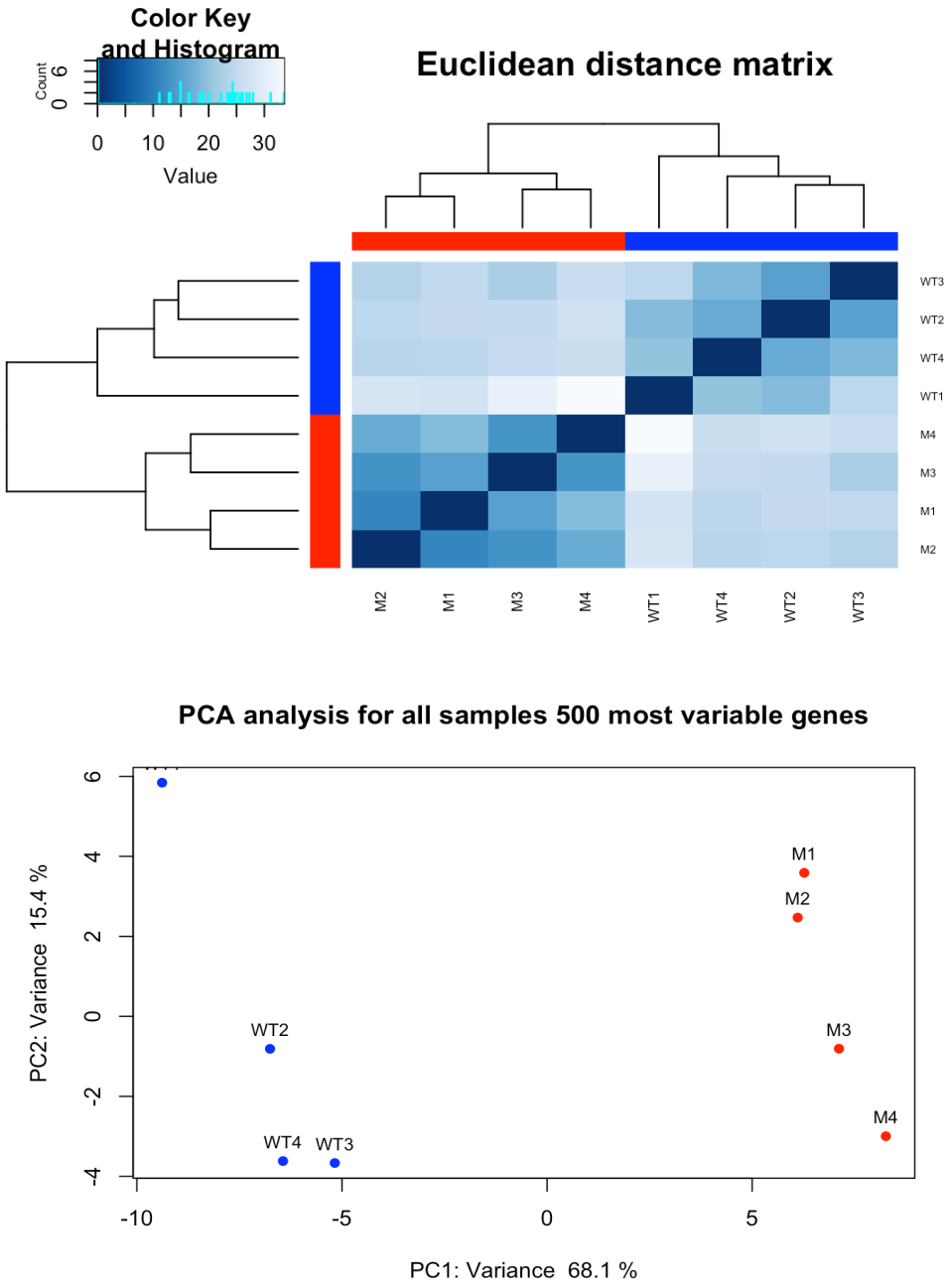


Figure 5-4 - Determining relatedness of replicates based upon gene expression profiles*. Top - Euclidean distance matrix visualising genetic distance and relatedness between replicates. The four wildtype replicates are depicted in blue, and the four rad21^{nz171} homozygote replicates, in red. Tree diagram shows the relatedness of the eight replicates. The alignment of the distance tree against itself demonstrates how similar gene expression is between groups, the darker the coloured box, the more similar the gene expression. Bottom - PCA plot, also demonstrating similarity between replicates. The four wildtype replicates are depicted in blue, and the four rad21^{nz171} homozygote replicates, in red.*

Both the distance matrix and the PCA plot analysis demonstrate that the four replicates for each condition are closely related, yet dissimilar to the other condition (Figure 5-4; further PCA plots in Appendix 4). The four wildtype replicates clustered together on the distance matrix tree, as did the four r*ad21*^{nz171} homozygote replicates. The two distinct groups clustered separately from one another. Self-alignment of the tree showed that gene expression profiles in wildtype replicates are similar to each other and dissimilar from *rad21*^{nz171} homozygote replicates, and vice versa. The PCA plot shows distinct clustering of the replicates for each condition, as well as separation between the wildtype and *rad21*^{nz171} homozygote samples, reinforcing that these replicates are suitable for comparative analyses.

5.5 Analysing differential gene expression

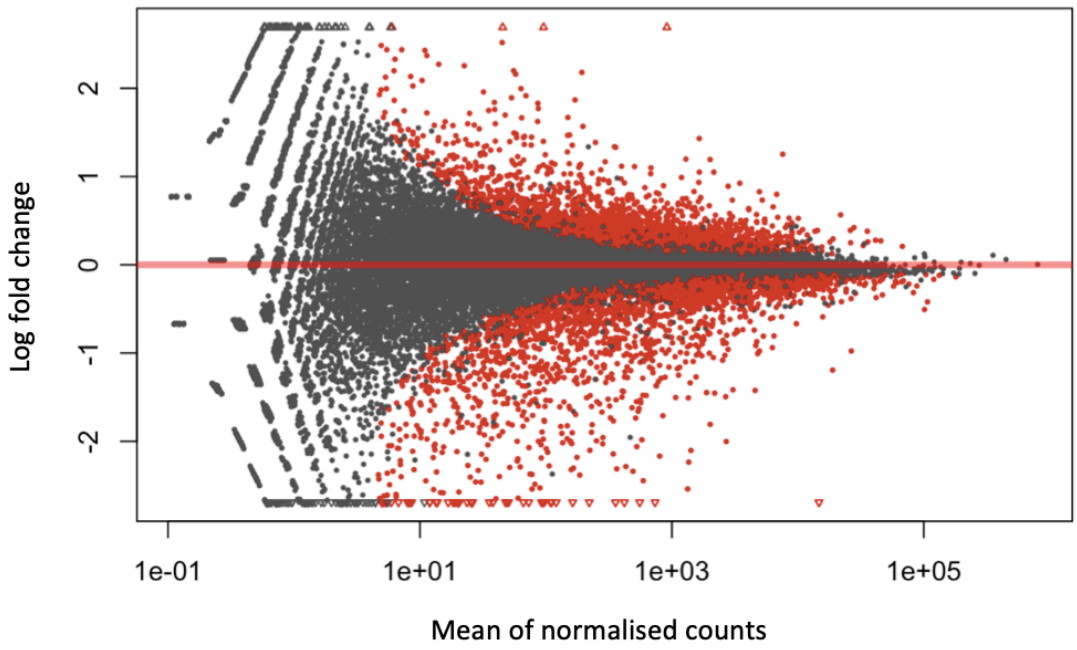


Figure 5-5 - MA plot of differential gene expression in rad21^{nz171} homozygote replicates compared to wildtype replicates*. The log fold change is plotted on the y axis, and the normalised expression signal is plotted on the x axis. All points on the plot are differentially expressed genes. Red points are significantly differentially expressed in rad21^{nz171} replicates compared to wildtype replicates. Corrected for multiple testing and false discovery rate.*

As the MA plot demonstrates, there were an abundance of genes significantly dysregulated in *rad21*^{nz171} replicates compared to wildtype (red points in figure 5-5). After applying statistical testing, there are 1988 genes that are up-regulated, and 2332 genes that are downregulated in *rad21*^{nz171} homozygote replicates compared to wildtype. This abundance of identified genes highlights how important cohesin is for normal gene expression and development.

It is expected to observe differential gene expression across the genome, which occurs by random chance, but statistical analysis of gene dysregulation per chromosome showed a significant increase on chromosome 12 and 16 (Appendix 4). Interestingly, the *rad21* gene is on zebrafish chromosome 16.

A number of the genes that were identified in these analyses had also been identified in *rad21*-deficient embryos at an earlier time point by Meier et al., warranting closer investigation. By analysing how these genes are expressed, by *in situ* hybridisation, quantitative data can be merged with visual data, allowing a better understanding of how these genes are dysregulated in a cohesin-deficiency. Both *rad21*-deficient and *stag1b*-deficient zebrafish embryos were assayed via *in situ* hybridisation. By comparing deficiencies in two different cohesin complex subunits, we aimed to examine if gene expression phenocopying occurs, or if the deficiencies in different subunits produce different expression patterns altogether. Heterozygous lines of *rad21*^{nz171} and *stag1b*^{nz205} fish were in-crossed to produce wildtype, heterozygous and homozygous mutant sibling offspring for fair gene expression comparison of *wnt3a, sema3ab, has2 and pax2a.* All four of these genes were identified as significantly dysregulated in the RNA-sequencing analyses. *Wnt3a* was chosen to examine because it is a master regulator of axial elongation and *sema3ab*, *has2* and *pax2a* have been identified in previous research.

5.5.1 Expression of *wnt3a* gene

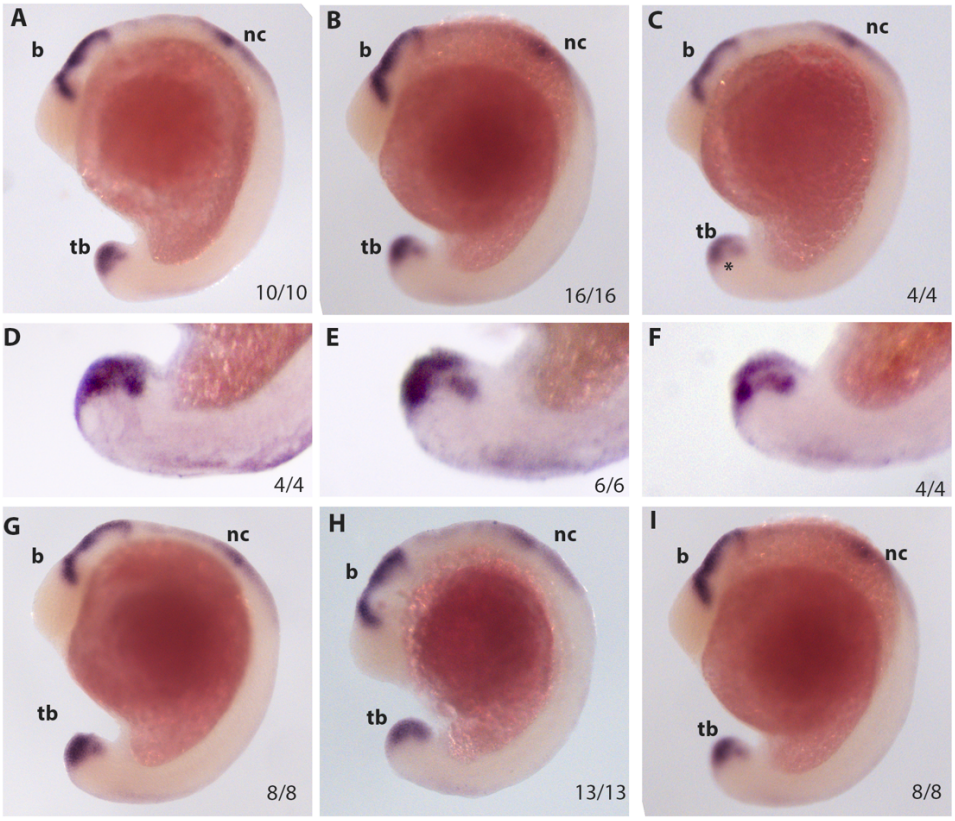


***Figure 5-6 - Expression of wnt3a in wild type, heterozygote and homozygote rad21^{nz171} and Stag1b^{nz205} embryos at the 18-somite stage**. A. Wildtype B. rad21^{nz171} heterozygote C. rad21^{nz171} homozygote D. Wildtype tailbud zoom E. rad21^{nz171} heterozygote tailbud zoom F. rad21^{nz171} homozygote tailbud zoom G. Wildtype H. stag1b^{nz205} heterozygote I. stag1b^{nz205} homozygote. Brain structures = b, neural crest cells = nc, tailbud = tb, * indicates a potential reduction in expression.*

Wnt3a was identified as significantly downregulated in the tailbuds of *rad21^{nz171}* homozygote embryos compared to their wildtype siblings in the RNA-sequencing analysis. The *in situ* gene expression patterns were not significantly different between Wild type, heterozygote and homozygote *rad21^{nz171} and stag1^{nz205}* embryos, as they all expressed *wnt3a* in the tailbud, the developing brain, and neural crest cells (Figure 5-6). *Wnt3a* tailbud expression was analysed at a higher magnification in *rad21^{nz171}* embryos, as whole embryo images (Figure 5-6 – A,B,C) indicated that there may be differential levels of *wnt3a*

tailbud expression between the genotypes. The star in figure 5-6 – C indicates a potential reduction in *wnt3a* expression in the tailbud of Rad21-deficient embryos, compared to those with functional cohesin. A second *in situ* hybridisation analysis was performed on another *rad21*^{nz171} heterozygous in-cross, and the tailbuds examined more closely. There is an observable decrease in tailbud *wnt3a* expression when an embryo is Rad21-deficient, *rad21*^{nz171} homozygote embryos (Figure 5-6-F) have patchier and diminished *wnt3a* levels compared to wildtype (Figure 5-6-D).

5.5.2 Expression of *sema3ab* gene

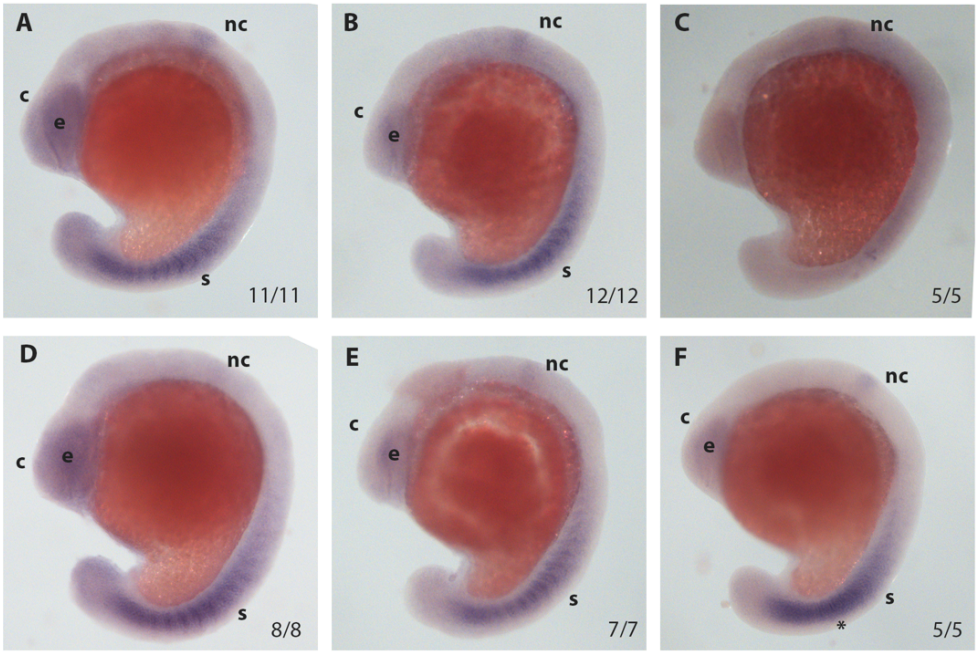


Figure 5-7 - Expression of sema3ab in wild type, heterozygote and homozygote rad21^{nz171} and stag1b^{nz205} embryos at the 18-somite stage. *A. Wildtype B. rad21^{nz171} heterozygote C. rad21^{nz171} homozygote D. Wildtype E. stag1b^{nz205} heterozygote F. stag1b^{nz205} homozygote. Craniofacial structures = c, neural crest cells = nc, somites = s, eye = e. * indicates an increase in expression in posterior somites.*

Sema3ab was identified as one of twelve semaphorin signalling molecules significantly downregulated in *rad21*^{nz171} homozygote tailbuds compared to tailbuds of their wildtype siblings, while another six semaphorin molecules were significantly up-regulated, making these genes crucial to examine spatially. The *in situ* gene expression patterns showed that

sema3ab is expressed in the craniofacial region, eye, somites and neural crest cells of wildtype and heterozygote *rad21^{nz171} and Stag1b^{nz205}* embryos (Figure 5-7 – A,B,D,E), making these genotypes almost indistinguishable by their *sema3ab* expression. Both *stag1b^{nz205}* and *rad21^{nz171}* heterozygotes appeared to have a reduction in *sema3ab* expression in the craniofacial region compared to their wildtype siblings yet retained expression in the eye. *Rad21*^{nz171} homozygote embryos had no *sema3ab* expression in the somites or the eye but retained expression in the neural crest cells (Figure 5-7 – C). *Stag1b^{nz205}* homozygote embryos had reduced *sema3ab* expression in the eye but retained expression in the neural crest cells and had increased expression in the somites, indicated by the * in Figure 5-7. Homozygous *rad21^{nz171} and stag1b^{nz205}* embryos were easily distinguishable from their siblings by *sema3ab* expression. *Sema3ab* signalling is important to guide neural crest cells into the pharyngeal arches, which will go on to form the upper and lower jaw, the face and palate (Yu & Moens, 2005), which explains the wide distribution of *sema3ab* around the eye and brain of the developing embryos. The semaphorin family of signalling molecules guide migration of a number of cell type, which could explain the expression seen in posterior somites (Nakamura et al., 2000).

5.5.3 Expression of *has2* gene

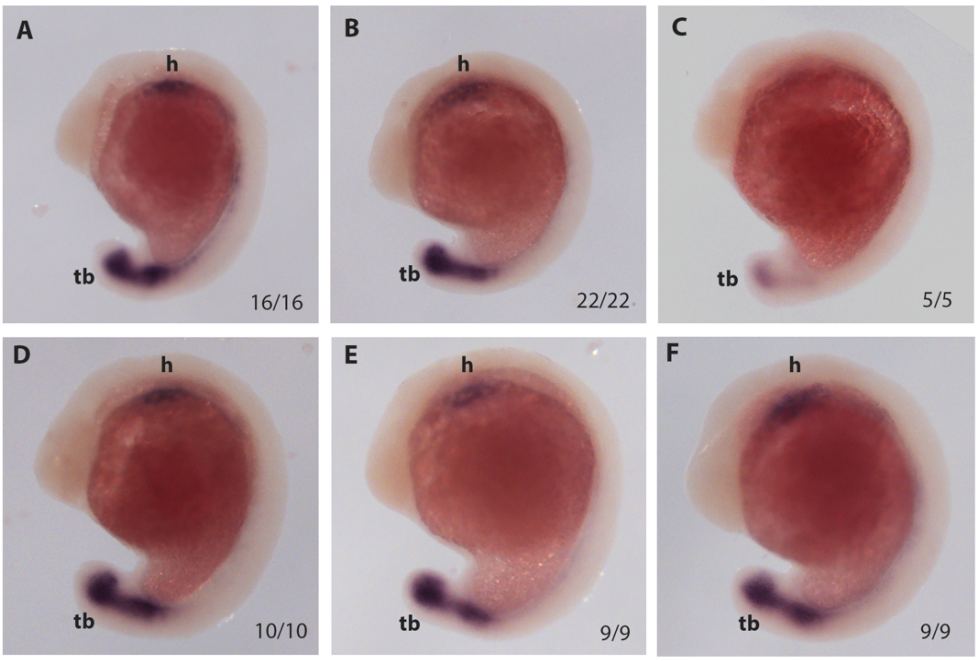


Figure 5-8 - Expression of has2 in wild type, heterozygote and homozygote rad21^{nz171} and stag1b^{nz205} embryos at the 18-somite stage*. A. Wildtype B. rad21^{nz171} heterozygote C. rad21^{nz171} homozygote D. Wildtype E. stag1b^{nz205} heterozygote F. stag1b^{nz205} homozygote. Heart = h, tailbud = tb.*

Has2 was identified as significantly downregulated in the tailbuds of *rad21^{nz171}* homozygote embryos compared to their wildtype siblings, *Has2* was expressed in the tailbud and the heart of wild type, heterozygote and homozygote *stag1b^{nz205}* embryos (Figure 5-8 – D,E,F) and wild type and heterozygote *rad21^{nz171}* embryos (Figure 5-8 – A,B), making these genotypes indistinguishable by their *has2* expression. *Rad21*^{nz171} homozygote embryos had diminished *has2* expression in the tailbud, and no expression in the heart, making them easily distinguishable from their siblings by *has2* expression (Figure 5-8 – C). *Has2* is essential for producing hyaluronan, a critical component of extracellular matrices and a major component of the cardiac jelly, which is critical for valve development (Just et al., 2016). In *rad21^{nz171}* homozygote embryos all heart *has2* expression was lost, consistent with improper heart valve formation (Schuster et al., 2015). The reduction in tailbud *has2*

expression of *rad21*^{nz171} homozygotes also indicates a breakdown of other essential structures which require strong extracellular matrices.

5.5.4 Expression of *pax2a* gene

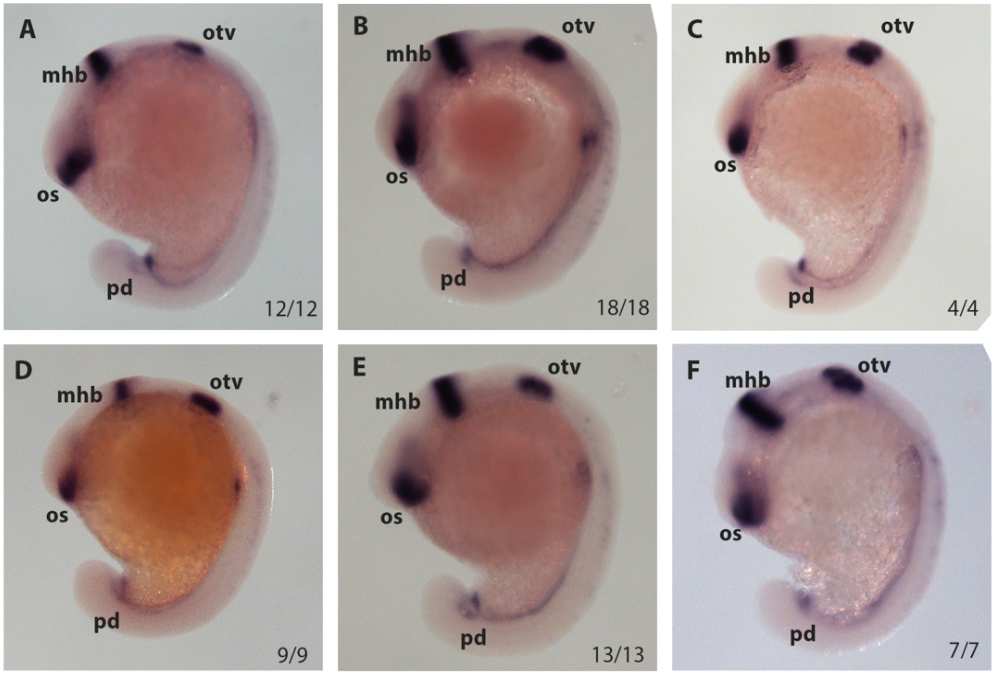


Figure 5-9 - Expression of pax2a in wild type, heterozygote and homozygote rad21^{nz171} and stag1b^{nz205} embryos at the 18-somite stage*. A. Wildtype B. rad21^{nz171} heterozygote C. rad21^{nz171} homozygote D. Wildtype E. stag1b^{nz205} heterozygote F. stag1b^{nz205} homozygote. Midbrain-hindbrain boundary = mhb, optic stalk = os, otic vesicle = otv, pronephric duct = pd.*

Pax2a was identified as significantly upregulated in the tailbuds of *rad21*^{nz171} homozygote embryos compared to their wildtype siblings. *Pax2a* was expressed in the developing eyes, brain, heart, spinal cord neurons and pronephric ducts of wild type, heterozygote and homozygote *stag1b*^{nz205} and *rad21*^{nz171} embryos (Figure 5-9), making these genotypes indistinguishable by their *pax2a* expression. *Pax2a* is a transcription factor involved in the development of a number of organs throughout embryogenesis, including the eye, kidney and the nervous system, which explains strong expression identified in these structures .

Though no differences were observed at this time point, it is established that *pax2a* gains expression in the tailbud of *rad21*^{nz171} homozygote embryos at around 24 hpf (Tahsin, 2019).

5.6 Pathway analyses

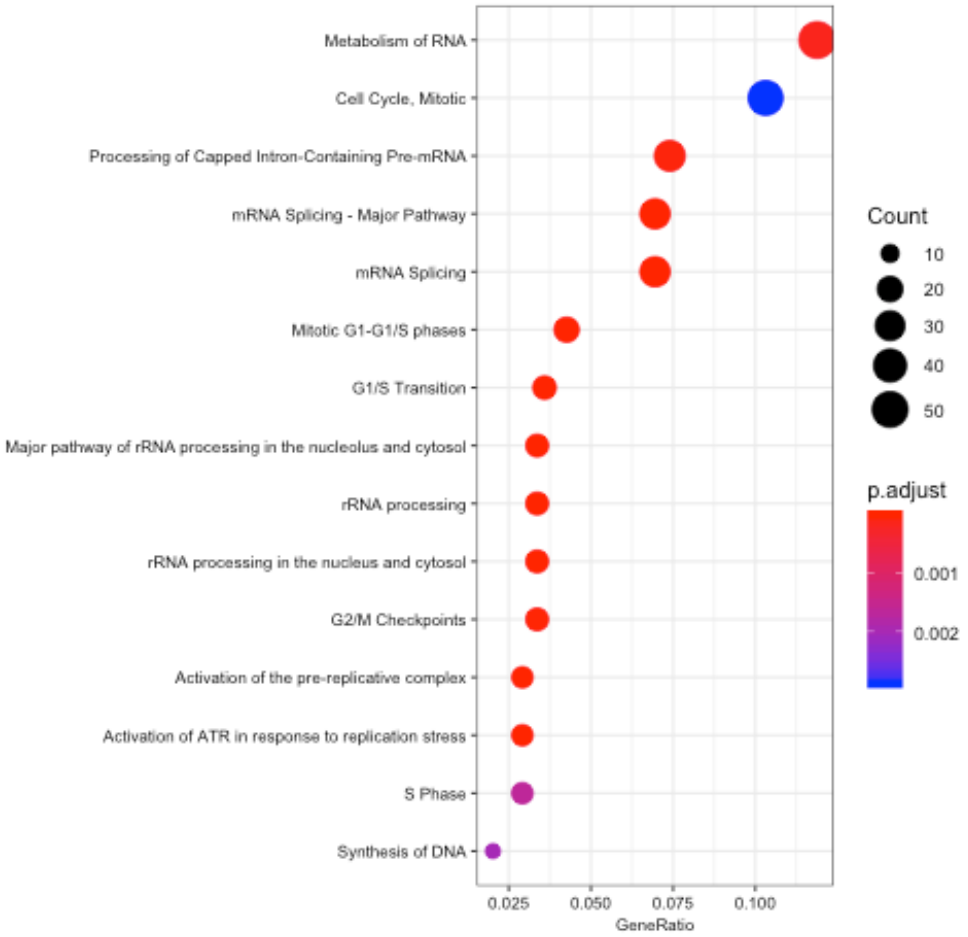


Figure 5-10 - Reactome pathway plot depicting the top 15 enriched pathways within the down regulated genes in rad21nz171 homozygote samples.

No biological pathways were enriched within the up-regulated genes in *rad21*^{nz171} homozygote samples. In comparison, 25 pathways were significantly enriched in down-regulated genes, with the top 15 depicted in figure 5-10. RNA processing was identified as extremely dysregulated in *rad21*^{nz171} homozygote samples. Many of the small nuclear ribonucleoprotein (snRNP) and pre mRNA processing factor (prpf) genes, which control mRNA splice site recognition and regulate alternative mRNA splicing were significantly downregulated, along with *nup107*, a nucleoporin gene, which allows mRNA transport from the nucleus. The cell cycle was also identified as extremely dysregulated in *rad21*^{nz171} homozygote samples, with key cell cycle regulators such as *cdk1* amongst a number of cyclin genes significantly downregulated in these samples. *rgcc*, the regulator of the cell

cycle, was also significantly downregulated in *rad21*^{nz171} homozygote samples. *rgcc* increases the kinase activity of the cyclin proteins, allowing progression of the cell cycle. Five of the six genes that make up the minichromosome maintenance protein complex (MCM) were significantly downregulated, and thus MCMs role in cell cycle DNA replication is likely deficient in *rad21*^{nz171} homozygotes. The reactome analyses are a database of reactions, pathways and biological processes, they are not cell type specific, or organ specific pathways, so these annotations must be elucidated elsewhere.

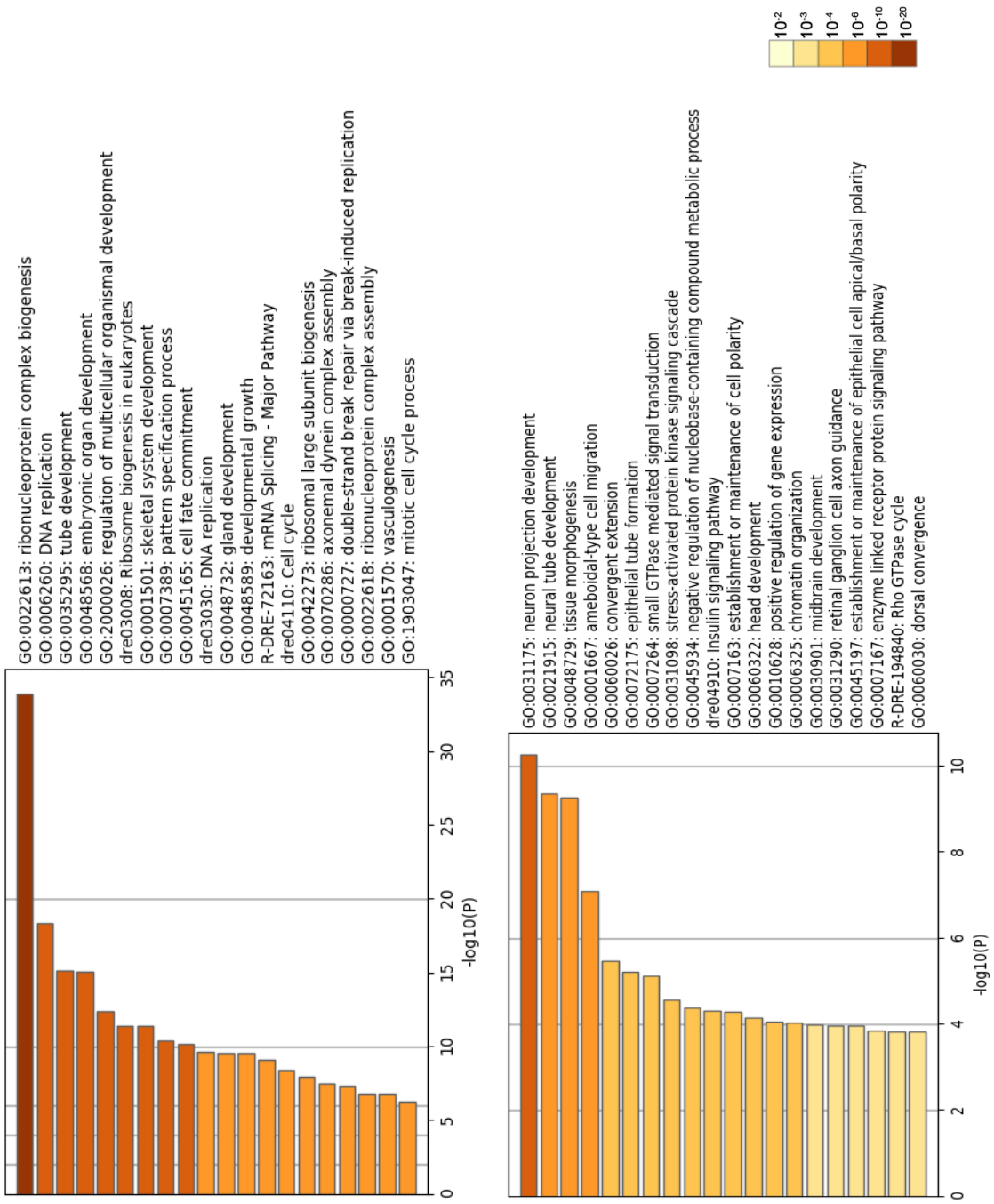


Figure 5-11 - Metascape analyses of significantly dysregulated pathways in rad21nz171 homozygotes. *Top: Metascape pathway enrichment for genes up-regulated in rad21^{nz171} homozygote samples compared to wildtype. Bottom: Metascape pathway enrichment for genes down-regulated in rad21^{nz171} homozygote samples compared to wildtype. The darker the bar, the more significant the result, as per scale.*

Metascape analyses group genes into function-relevant annotations, and identified neural system development pathways to be the most significantly upregulated in *rad21*^{nz171} homozygotes compared to wildtype. A large number of genes associated with neural tube formation and closure were upregulated, such as *ccdc85b, scrib, pcdh19* and *neo1a*. *Nrp2b*, a binding factor for the semaphorin family was upregulated, along with a number of other signalling molecules crucial for axon guidance and neural crest cell migration.

A varied list of pathways were just as significantly downregulated, including DNA and RNA processing and cell cycle control, as identified in Reactome analyses. Interestingly, Metascape identified ribosome biogenesis as the most significantly downregulated pathway in *rad21*^{nz171} homozygotes. RNA processing and ribosome biogenesis are not currently well-established roles for cohesin within a cell, and warrant further examination. Along with the genes identified as key in the processing of pre-mRNA, genes involved in formation of the ribosome, which will translate the mRNA into functional protein, were also significantly downregulated. The ribosome protein large (*rpl*) genes were abundantly downregulated, as were the dead box (*ddx*) genes, which arrange ribonucleoprotein complexes. *nmd3* and *eif6*, two genes involved in exporting ribosomal subunits from the nucleus were also downregulated.

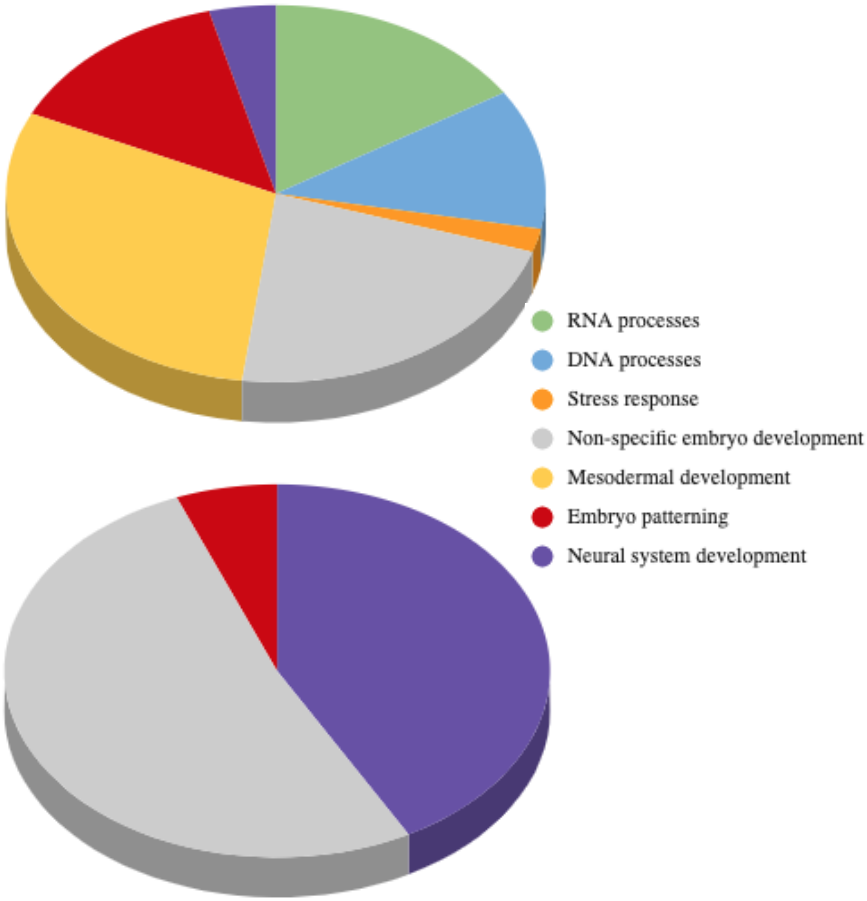


Figure 5-12 - Pie charts demonstrating the breakdown of pathway identities in the top 50 up and down regulated pathways enriched in rad21nz171 homozygote samples compared to wildtype*. Top – Downregulated pathways. Bottom – Upregulated pathways*

Of the top 50 most significantly up-regulated pathways in *rad21*^{nz171} homozygote samples, most were non-specific, but a significant portion were related to neural system development, while another portion were related to embryo patterning. In comparison, of the 50 most significantly down-regulated pathways, many more specific processes were affected, rather than generalised developmental pathways. Most significantly, RNA and DNA processing pathways were downregulated, as were pathways related to forming mesodermal systems. Embryo patterning pathways were also downregulated. Together these data indicate that when an embryo is Rad21 deficient, on top of expected dysregulation in cell cycle pathways and DNA processing, there is a shift from mesodermal to neural tissue production.

Previous work by Farrell et al involved single cell sequencing of zebrafish embryos, and applying a computational approach to describe the transcriptional profiles of a number of distinct cell trajectories. This work was carried out in zebrafish embryos from 3 hours post fertilisation to 11.5 hours post fertilisation (Farrell et al., 2018). Though our analyses examined gene expression profiles at 18 hours post fertilisation, the dysregulated gene lists were compared to Farrell trajectory lists, to determine if any general trends in trajectory dysregulation were still relevant. A large number of the genes identified in these analyses overlapped with the key genes for cell trajectories identified in Farrell et al (Appendix 4). Most significant was the finding that a number of tailbud genes that Farrell et al associated with a mesodermal fate were downregulated in our analyses, while tailbud genes that Farrell et al associated with a neural fate were upregulated in our analyses (Appendix 4). Many of the genes Farrell et al found to be key in underlying haematopoietic pronephron and cephalic mesoderm trajectories, two mesodermal fates, were significantly downregulated in our analyses (Appendix 4), further reinforcing the finding there is a skew from mesodermal fate to neural fate in *rad21*^{nz171} homozygote embryos.

5.7 Conclusion

The RNA sequencing analyses provided an abundance of data for further exploration. 2332 genes were significantly downregulated in cohesin-deficient embryos and 1988 genes were significantly upregulated, highlighting just how crucial cohesin is in regulating gene expression. The *in situ* analyses on a number of the genes identified both in these and previous analyses highlighted in a visual manner how essential cohesin is for gene expression, and that deficiency in different cohesin subunits does not produce the same gene expression changes, likely due to the redundancy of some proteins compared to others. In our analyses the ability of the Stag proteins to compensate for one another was evident, yet when examining *sema3ab* expression *stag1b*^{nz205} homozygotes had an increase in expression in the posterior somites which indicates an affinity for *stag1b*-bound cohesin in regulating this gene. The similarity in expression patterns between *rad21*^{nz171} heterozygote and wildtype siblings indicates that one functional copy of *rad21* is sufficient for regulating gene expression. *Rad21*^{nz171} homozygote embryos, with no functional cohesin complexes formed, had significant changes in *has2* and *sema3ab* gene expression

patterns. *Pax2a* expression patterns appeared consistent across all three genotypes from a *rad21*^{nz171} heterozygote in-cross, phenocopying the results from a *stag1b*^{nz205} in-cross. *Wnt3a* expression patterns in *rad21*^{nz171} embryos also phenocopied those seen in *stag1b*^{nz205} embryos, despite *wnt3a* expression in the tailbud of *rad21*^{nz171} homozygotes appearing diminished in comparison to their wildtype siblings.

The pathway analyses provided numerous avenues for further investigation. Due to cohesins established role in cellular processes such as mitosis and DNA replication, it is not surprising these were identified were significantly downregulated. Of interest is the apparent skew to neural ectoderm, as well as a potential role for cohesin in regulating ribosome biogenesis and therefore, protein translation.

6 THE UNCONSTRAINED CHROMATIN THEORY

Cohesin is necessary for the regulation of many genes, but it is not clear how cohesin may be controlling their expression. We predict gene dysregulation to be a product of unconstrained chromatin in a cohesin-deficiency allowing ectopic signals to be interpreted. We have found that a major consequence of cohesin mutation is that cells skew away from mesoderm commitment and towards neuroectoderm in zebrafish tailbuds. Commitment to mesoderm is controlled by a Wnt signalling gradient in the tailbud (Martin et al., 2012). We hypothesised that disrupted Wnt signalling is partially responsible for cell fate changes in the tailbud, as indicated by a downregulation of Wnt3a (Figure 5-6). The zebrafish Wnt reporter line Tg(7xTCFXla.Siam:nlsmCherry)ia5 line expresses the mCherry fluorescent protein when Wnt target genes are activated. Stimulations with BIO-6 (a GSK-3 inhibitor) should allow more β-catenin to translocate to the nucleus, switching on mCherry. We hypothesise that in cohesin-deficient embryos, we will observe a difference in Wnt signal transduction. We predict that the levels of mCherry expression will differ in cohesin-deficient embryos compared to wildtype. Comparing mCherry expression in cohesin-deficient fish to mCherry expression in wildtype fish acts as a model for dysregulated Wnt signalling as a result of cohesin deficiency and is the first step in exploring the unconstrained chromatin theory.

6.1 Toxicity testing of the WNT agonist

The appropriate dosage for BIO (2'Z,3'E)-6-Bromoindirubin-3'-oxime (BIO-6) stimulation was determined by conducting maximum tolerated dose studies. BIO-6 was tested at 6 concentrations, 1, 2.5, 5, 7.5 and 10 µM, with doses higher than 5 µM being too toxic and 1 µM being inefficacious. 2.5 µM was chosen as the effective treatment dosage, as embryos survived, and appeared developmentally normal compared to the 5 µM dose (Appendix 4).

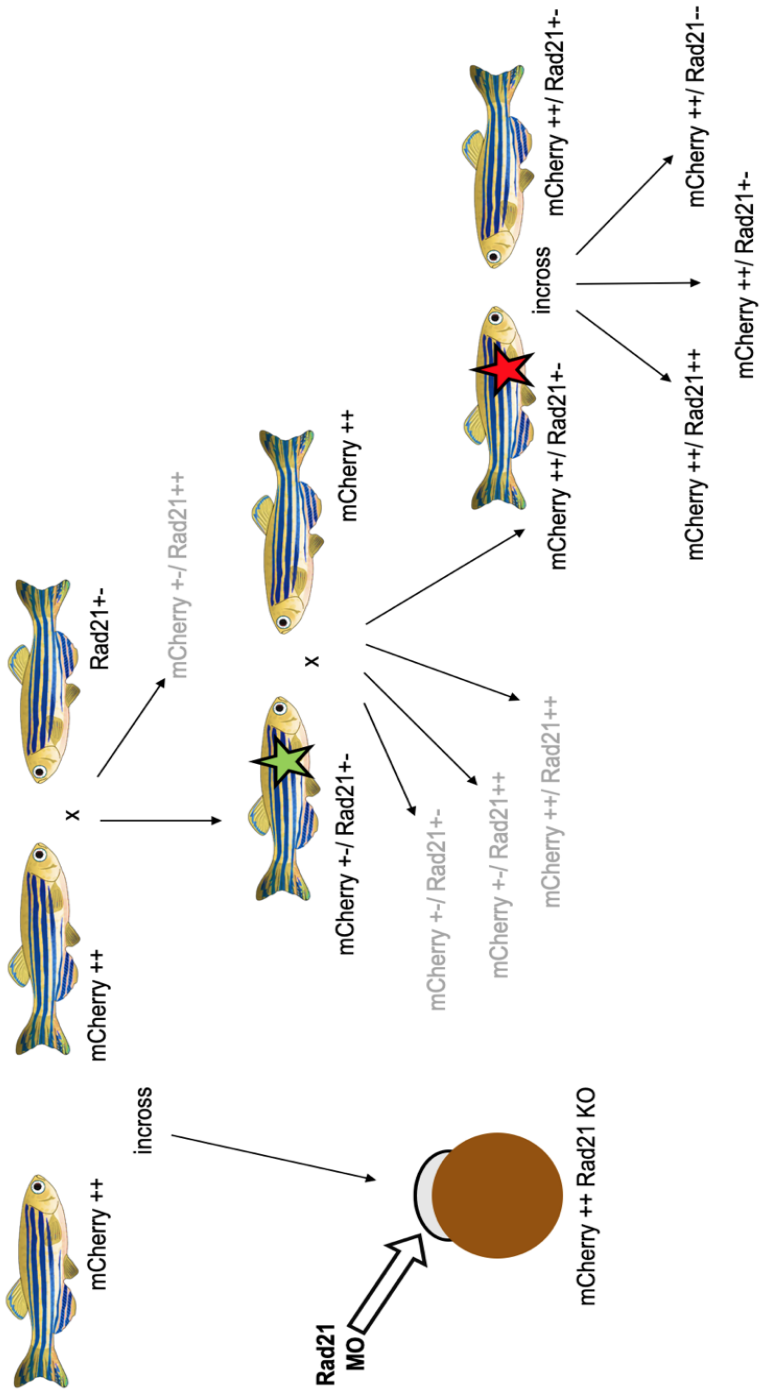


Figure 6-1 - Fish experiments to examine the effect of cohesin-deficiency on Wnt signalling*. For simplicity, genotypes denoted as ++ for homozygous functional, +- for heterozygous and -- for homozygous mutant. Tg(7xTCFXla.Siam:nlsmCherry)ia5 fish referred to as mCherry in the diagram. Left: Rad21 morpholino injections into Tg(7xTCFXla.Siam:nlsmCherry)ia5 fish incross. Right: Outline of the necessary fish crosses required to generate homozygous mCherry-expressing embryos with comparable cohesin deficiencies.*

6.2 Cohesin knockdown with rad21 morpholino

A morpholino binds to endogenously produced mRNA, blocking its translation and intended function. Previous injections with a *rad21* morpholino have been successful in knocking down *rad21* function to explore the proteins normal role within a cell (Horsfield et al., 2007). Injections of *rad21* morpholino into Tg(7xTCFXla.Siam:nlsmCherry)ia5 embryos, and stimulation with BIO-6 resulted in embryo death, despite the previous success of the morpholino at these concentrations (Meier et al., 2018) and BIO-6 toxicity screening indicating that embryos can tolerate this dosage. Reducing the concentration of morpholino any further would not result in sufficient cohesin knockdown so that "knockdown" embryos could be compared to wildtype, and reducing the BIO-6 concentration was not supported in the literature. Creating a stable line of transgenic, cohesin mutant fish was necessary to examine the effect of cohesin deficiency on WNT signalling.

6.3 Creating a stable line of transgene carrying, cohesin deficient fish

Figure 6-1 outlines the workflow necessary to create the Tg(7xTCFXla.Siam:nlsmCherry)ia5 homozygote/*Rad21*^{nz171} heterozygote fish line required for further experiments. *rad21*^{nz171} heterozygote fish were crossed to Tg(7xTCFXla.Siam:nlsmCherry)ia5 fish, and the offspring were grown out to adults, where they were TaqMan genotyped at the *rad21*^{nz171}locus. Only Tg(7xTCFXla.Siam:nlsmCherry)ia5 heterozygote/ *rad21*^{nz171} heterozygote fish were kept for further crosses (Green star in figure 6-1).

Tg(7xTCFXla.Siam:nlsmCherry)ia5 heterozygote/ *rad21*^{nz171} heterozygote fish were in-crossed, and the embryos examined with confocal microscopy at 18 hpf (Appendix 4). Clear differences in the amount of fluorescence was evident between embryos, due to the offspring from this cross being either homozygous or heterozygous for Tg(7xTCFXla.Siam:nlsmCherry)ia5 genotypes (See offspring genotypes in figure 6-1). It became apparent that stimulating fish from these in-crosses would not produce quantifiable changes in mCherry expression between *Rad21* genotypes, as the embryos also had different degrees of mCherry expression based upon the number of copies of the

transgene inherited. To assay whether BIO-6 was toxic to mutant embryos, stimulations on a batch of the in-crosssed embryos was performed. The embryos survived, confirming that the 2.5 µM dosage is suitable.

A back cross of Tg(7xTCFXla.Siam:nlsmCherry)ia5 heterozygote/ *rad21*^{nz171} heterozygote fish to the parent Tg(7xTCFXla.Siam:nlsmCherry)ia5 line is necessary, to produce offspring with two copies of the transgene. As fluorescence levels indicate how many copies of the transgene an embryo is carrying (Appendix 4), homozygous Tg(7xTCFXla.Siam:nlsmCherry)ia5 embryos should be selected to be grown out to adult, where a fin clip can be taken and the fish genotyped for their *rad21*^{nz171}status, with only Tg(7xTCFXla.Siam:nlsmCherry)ia5 homozygote/ *rad21*^{nz171}heterozygote fish kept for further experiments (Red star in figure 6-1). Tg(7xTCFXla.Siam:nlsmCherry)ia5 homozygote/ *rad21*^{nz171} heterozygote fish can be in-crossed to produce embryos that all have two copies of the transgene, but different functional *rad21* dosages, making quantifiable mCherry expression changes possible across cohesin-deficient genotypes. Due to time constraints, the backcross and subsequent genotyping has not been performed.

6.4 Conclusion

Knocking down *rad21* expression in Tg(7xTCFXla.Siam:nlsmCherry)ia5 embryos and stimulating with BIO-6 was toxic to the embryos, yet stimulating Tg(7xTCFXla.Siam:nlsmCherry)ia5 heterozygote/ *rad21*^{nz171} heterozygote in-crossed embryos with 2.5 µm BIO-6 was not toxic to the embryos, indicating that the dosage is appropriate in mutant embryos, but not in combination with morpholino knockdown. Further fish crosses and genotyping must be performed so that stable lines of Tg(7xTCFXla.Siam:nlsmCherry)ia5 homozygote/ *rad21*^{nz171} heterozygote fish are available for in-crossing and subsequent BIO-6 stimulations to allow for the analysis of the Wnt signalling pathway in Rad21-deficient embryos.

7 DISCUSSION AND FUTURE DIRECTIONS

7.1 Examining RAD21 functionality with a *runx1* bioassay

The zebrafish *runx1* bioassay determined that the familial variant S449G-RAD21 is functional and the novel I152F-RAD21 variant is non-functional. It is not surprising that the familial variant was found to be functional, as the cohort of patients do not present with any cohesinopathy spectrum phenotypes, as would be expected if there was a reduction in functional cohesin in these patients (Horsfield et al., 2012). Though this bioassay confirmed that S449G-RAD21 is functional, it cannot determine if there are slight changes in the functionality of the cohesin complexes it forms. Knowing that S449G-RAD21 is functional shifts the focus from mitotic-failure driven cancer progression to gene expression-driven cancer and the way in which S449G-RAD21 functions in human cells to regulate gene expression must be explored. It may be that S449G-RAD21 can form cohesin complexes so that mitosis, DNA repair and gene expression can happen, but slight changes in cohesin binding or topological organisation of the chromatin results in cancer-driving gene expression changes. The novel variant I152F was produced at the opposite end of the *RAD21* gene to S449G and introduced a large phenol group to the protein, in a bid to create protein folding changes or steric hindrance at the binding interface with SMC3. The bioassay determined that the novel variant is non-functional, and protein analysis determined that the variant site has potential to interfere with the binding to the SMC3 ATPase head. It is possible that despite not being directly involved in binding the SMC3, that the changes in I152F-RAD21 still affect cohesin complex formation or interactions with any of the other cohesin-binding regulators. It would be interesting to determine if this variant is identified in any future patients on the cohesinopathy spectrum.

As there is no full RAD21 protein structure available; the sites where the variants are predicted to be located are only approximate. When a full structure becomes available, re-examining the location of the variants within the 3D protein may shed further light onto the interactions that the I152F variant is involved in. The most significant outcome from these analyses was identifying how crucial functional annotation of variants is, especially in

a clinical setting. All of the protein prediction pipelines produced varying results about the pathogenicity of the two variants, both from one another and from our functional analyses. Though prediction algorithms are becoming increasingly more informative and accurate, they are not perfect, and when it comes to patient wellbeing, it is important to get them the most accurate information possible, through functional analysis.

7.2 The role of cohesin in normal zebrafish embryogenesis

The high-throughput RNA sequencing and analysis identified over four thousand significantly dysregulated genes in cohesin-deficiency. We examined just four of these via *in situ* hybridisation. *Wnt3a* was found to be significantly down-regulated in *rad21*^{nz171} homozygote tailbuds compared to wildtype and as this gene is one of the master regulators of axial extension, we wanted to examine its expression pattern in wildtype and *rad21*^{nz171} tailbuds. We chose the other three genes because they overlap with genes previously identified as dysregulated in cohesin-deficiency at an earlier time point in embryogenesis (Meier et al., 2018). Assaying the expression patterns of the identified genes is necessary to build a bigger picture of cohesin-deficient embryo development. We hoped to visualise if any of the four genes exhibited marked differences in expression patterns on embryos produced from the in-cross of adult heterozygote *rad21*^{nz171} and *stag1b*^{nz205} zebrafish. In this manner, any dysregulation could be examined across wildtype, heterozygous and homozygous mutant siblings, as well as between different cohesin-subunit genotypes.

7.2.1 Comparing gene expression patterns across cohesin subunits

As there are four Stag orthologues in zebrafish, expression patterns of the four genes were expected to be consistent across the different *stag1b*^{nz205} genotypes, due to any of the other *stag* orthologues compensating for a loss of *stag1b*. Compensation was apparent for regulating all genes except for *sema3ab*, in which the *stag1b*^{nz205} heterozygotes had a decrease in craniofacial expression, and homozygotes had distinctly different expression in the somites, indicating a favourability of *stag1b*-associated cohesin in regulating *sema3ab* expression. As there was not a total reduction in *sema3ab* expression, it is clear that orthologous *stag* proteins are sufficient to support regulation, also supported by the fact

that a total *stag1b* knockout is survivable (Muto & Schilling, 2017). As *stag1b*^{nz205} fish are homozygous viable, serious posterior patterning defects were not expected.

Unlike a homozygous *stag1b* loss, a homozygous *rad21* loss is lethal, as there are no compensatory proteins that can perform its role, and thus a homozygous *rad21* loss produces a total cohesin knockout (Sonoda et al., 2001). One of the key observations from the expression analyses was how indistinguishable *rad21*^{nz171} heterozygotes are from their wildtype siblings, despite having a 50% reduction in functional cohesin. This indicates that one functional copy of the *rad21* allele is sufficient for normal expression of most of the genes during development, and is supported by the fact that mutations in *rad21* tend to produce mild cohesinopathy phenotypes (Minor et al., 2014), as the functional *rad21* allele can compensate for the non-functional so that development is relatively unaffected.

We hoped to examine if a reduction in different cohesin subunits produced similar gene expression patterns. Similarities in spatiotemporal gene expression between *stag1b*^{nz205} and *rad21*^{nz171} heterozygotes and their wildtype siblings was clear and is evidently due to the fact that a single copy of *rad21* is sufficient for expression, and likewise a single copy of *stag1b*, in combination with numerous orthologous proteins, allows normal gene expression. *Pax2a* gene expression was similar for *rad21*^{nz171} and *stag1b*^{nz205} homozygous embryos.

Stag orthologues were sufficient to regulate *wnt3a* and *has2* gene expression in the absence of *stag1b,* producing homozygous mutants with phenotypes indistinguishable from their siblings with functional *stag1b*. *Rad21*^{nz171} homozygous embryos have no orthologous protein redundancy pathways available, and *wnt3a* and *has2* gene expression patterns were consequently affected. Interestingly, *stag1b*^{nz205} homozygous embryos did not have similar *sema3ab* expression patterns to their siblings with functional *stag1b*, and the change in expression pattern did not phenocopy the total loss of *sema3ab* expression observed in *rad21*^{nz171} homozygous embryos, indicating that while a total *stag1b* loss affects gene expression, it is not to the same degree that a *rad21* loss does. These findings

are consistent with the fact that a total *rad21* knockout is lethal, while a *stag1b* knockout is survivable.

It is important to note that though differences in expression patterns were not strikingly obvious at this time point for all four genes, it does not mean that at other points in embryogenesis these genes are not significantly visibly different in cohesin-deficient embryos compared to wildtype. As all four of these genes were identified as significantly dysregulated in cohesin-deficient embryos, there may be very small differences in gene expression patterns across these genotypes that could not be quantified, and that in reality may underlie some of the significant clinical presentations of the cohesinopathies.

7.2.2 Unsurprising dysregulated pathways in cohesin knockout

As cohesin is so fundamental in key cellular processes, it is no surprise to see a large number of these significantly downregulated in the pathway analyses. DNA replication, cell cycle, double strand break repair and mitotic control pathways were all significantly down-regulated, consistent with previous research that establishes cohesin as a regulator of these processes. Downstream of these key biological processes, organ development, embryo patterning and general growth pathways are consequently also significantly downregulated.

7.2.3 A role for cohesin in regulating gene expression through translation

Interestingly, a number of the most down-regulated processes were those relating to ribosome biogenesis and ribonucleoprotein complex biogenesis, yet there is limited research supporting a role for cohesin in regulating gene expression through enabling RNA translation and protein production. Bose et al show that in yeast and human cells, Roberts Syndrome associated mutations, in SMC and ESCO-1, a cohesin loading factor, less ribosomal RNA is produced and the effect of this is expected to constrain protein synthesis, thus essential proteins are not translated, and developmental defects arise (Bose et al., 2012). Our findings support a reduction in essential protein production machinery when cohesin is lost, which brings a whole new level of cohesin regulation on gene expression to the forefront. Not only can cohesin regulate gene expression at the transcriptional level

through chromatin organisation, it may be essential in helping the translation of these genes from mRNA to protein.

Ribosomes are so essential they are synthesised within a cell at a rate of thousands per minute (Fatica & Tollervey, 2002). Defects in ribosome biogenesis or ribonucleoprotein complex biogenesis would therefore be expected to be lethal, as without creating thousands of these structures, essential proteins cannot be formed effectively. A number of human diseases related to ribosome-biogenesis are surprisingly survivable, resulting in tissue-specific pathologies, a number of which have clinical presentations that overlap with the cohesinopathy spectrum. Mutations in TCOF, a ribosome biogenesis protein, results in craniofacial abnormalities similar to CdLS, with large eyes and small chins (Freed et al., 2010). Mutations in RMRP, a small nucleolar ribonucleoprotein, phenocopy CdLS also, with short stature and mental retardation (Freed et al., 2010). It is possible that some of the phenotypes observed in the cohesinopathy spectrum are due to issues in cohesins ability to assist in ribosome biogenesis.

7.2.4 Gene dysregulation is not equal across zebrafish chromosomes

An unexpected finding was that gene dysregulation was not equal across all zebrafish chromosomes. Chromosome 12 and 16 had significantly increased levels of dysregulation compared to all other chromosomes. Interestingly, *rad21* is located on zebrafish chromosome 16. It is possible that in the absence of functional *rad21*, cell transcription machinery preferentially transcribed from chromosome 16, in a bid to produce *rad21* proteins, and in doing so increased the expression of surrounding genes.

7.2.5 Stem cell fates skewed towards neural ectoderm

Perhaps the most striking finding was that bipotent stem cells in a cohesin-deficient tailbud appear to skew cell fate decisions towards neural ectoderm, as there was a significant increase in the expression of genes related to neural processes and neural pathway upregulation and a significant decrease in mesodermal processes and mesodermal pathway downregulation.

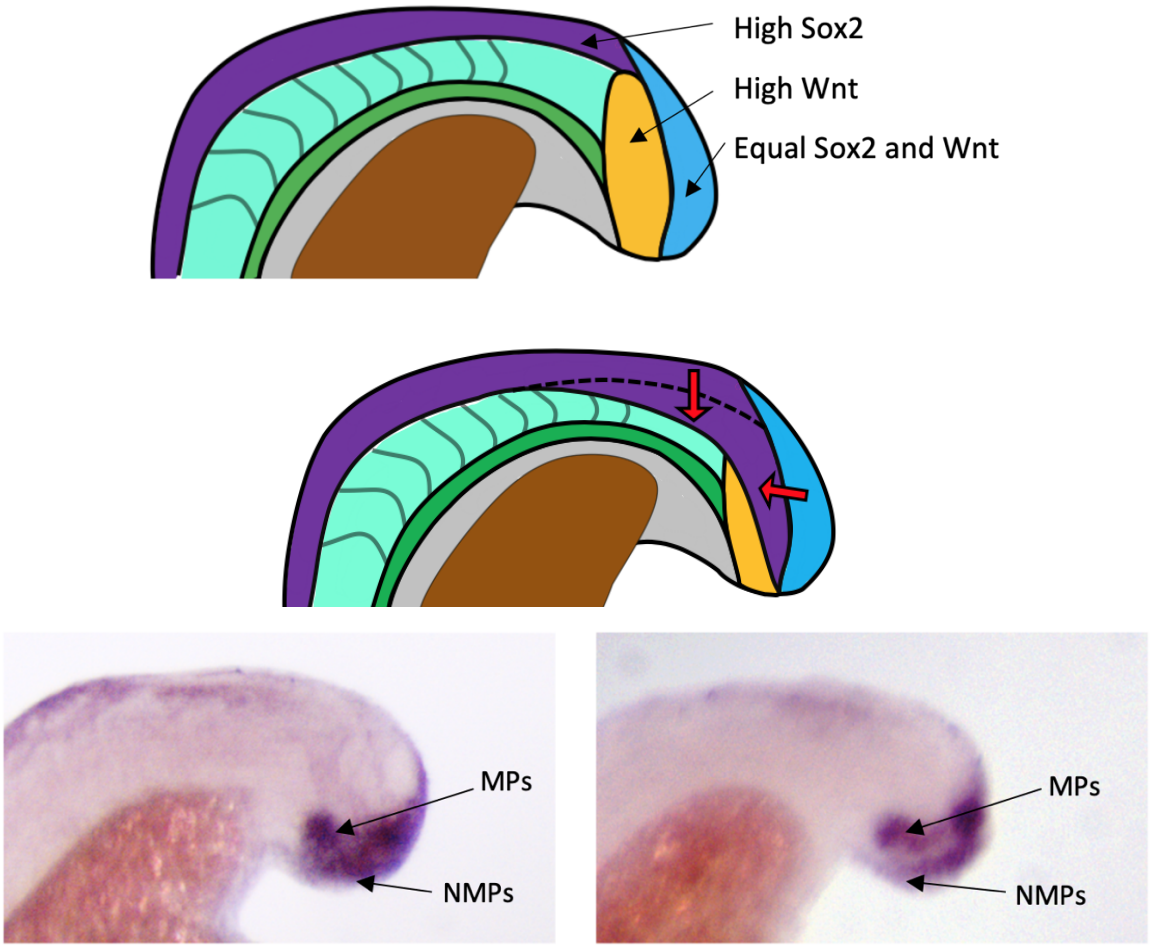


Figure 7-1 - Tailbud tissues, and proposed misspecification models. *Schematics of tailbud structures within the zebrafish tailbud. The tailbud is oriented A-P. Blue – Neuromesodermal progenitor population. Yellow – Mesodermal progenitor population. Purple – Neural ectoderm. Green – vascular endothelium. Teal – Presomitic mesoderm. Stripes in the presomitic mesoderm represent formed somites. Top tailbud schematic depicts the normal distribution of tissues in a developing tailbud and shows expression levels of key genes in the tissues. Middle tailbud schematic depicts potential expansion of neural-fated cells and two proposed ways in which cohesin-deficient embryos may gain this. The top red arrow represents a spatial expansion of neural progenitors, while the bottom red arrow represents NMPs specifying into neural cells rather than mesodermal progenitors. The bottom panel shows wnt3a expression in wildtype (left) and rad21^{nz171} homozygote (right) tailbuds at 18hpf. Images originally from figure 5-6 but rotated to allow comparison to tissue distribution schematics. Proposed sites for neuromesodermal progenitors (NMPs) and mesodermal progenitors (MPs) labelled.*

The first decision that a neuro-mesodermal progenitor stem cell (NMP) at the very posterior of the tailbud must make is between a neural or mesodermal fate. The schematics in figure 7-1 depict an elongating tailbud, and shows the distribution of these tissues throughout the structure. (Gouti et al., 2014). *Wnt3a* and *wnt8* signalling drives NMPs towards a mesodermal fate (yellow in figure 7-1), and a reduction in Wnt signaling drives a neural fate, as *sox2* levels increase (purple in figure 7-1) (Gouti et al., 2014). In RNA sequencing and subsequent *in situ* hybridisation analyses, we showed that *wnt3a* expression is downregulated in *rad21*^{nz171} tailbuds compared to wildtype, thus an increased number of progenitor cells are likely to specify into neural rather than mesodermal fates.

The NMPs arrest in G2 phase, just before mitosis, ahead of fate commitment (Mastromina et al., 2017). In *rad21*^{nz171} homozygotes, NMPs take longer to progress through mitosis as Rad21-deficient embryos cannot perform the cell cycle efficiently (Appendix 4). It may be that holding NMPs in G2 has some effect on cell fate decisions which contributes to the observed skew to neural ectoderm, though this remains to be explored.

Though *wnt3a* levels are significantly downregulated in *rad21*^{nz171} homozygotes, some NMPs will still have sufficient *wnt8* expression levels to specify into mesodermal progenitors, supported by the fact that these *rad21*^{nz171} homozygote embryos are phenotypically indistinguishable from their wildtype siblings at this stage, where there are numerous mesodermal structures such as the somites present. Up until this time point, the genotypically homozygous mutant embryo has not been a total cohesin knockout. Maternal constructs, identified in trace amounts in genotypically mutant embryos, could also be assisting in the production of mesodermal progenitors from NMPs. The exciting aspect of assaying embryo expression profiles at this stage in development is that this is the cusp of no residual maternal transcripts and therefore the effect of a total loss of cohesin on expression at this developmental stage can be observed.

There are two ways in which neural processes are appearing upregulated at the expense of mesodermal. Either NMPs are undergoing a cell fate decision change in the absence of sufficient Wnt signalling and becoming neural, or neural progenitors are produced as

normal, due to up-regulated *sox2* signalling, and because there are fewer mesodermal progenitors being produced, there is an expansion of neural progenitors into space that would normally be occupied by mesodermal progenitors (Figure 7-1 - middle).

7.2.6 Wnt signalling is required throughout axial elongation

Wnt signalling drives multiple decisions in axial elongation. FGF and Wnt normally upregulate *tbx6/16* in mesodermal progenitor cells (Yellow in figure 7-1), which suppresses *sox2*, and associated neural fate (Goto et al., 2017). A number of FGF genes were upregulated in the cohesin knockouts, and the Tbx genes remained unaffected. There is potential that the upregulation of FGF genes compensates for the downregulation of Wnt, maintaining *tbx6/16* expression, and thus maintaining a mesodermal progenitor state by suppressing *sox2*.

Those NMPs that do successfully differentiate into mesodermal progenitors require Wnt signalling for a second cell fate decision, between paraxial mesoderm and vascular endothelium (Teal and green respectively in figure 7-1) (Goto et al., 2017). High Wnt expression drives a number of genes required for transitioning into paraxial mesoderm, including *msgn1* – the master regulator of paraxial mesodermal fate (Yabe & Takada, 2012). Despite the upregulation of FGF, which is thought to supplement WNTs activation of this gene, *msgn1* is severely down-regulated, in fact it is the most significantly down-regulated gene identified in the analyses, likely because *wnt3a* cannot activate it. It is likely that the paraxial mesoderm and related structures will not form correctly in cohesin-deficient embryo from this time point, as there is now no maternal deposited RNA to supplement development. A low Wnt gradient in mesodermal progenitors allows cells to differentiate into vascular endothelium (Dunty et al., 2008). In our analyses there is no upregulation of vascular processes, rather a down-regulation. As it appears there has been a significant downregulation of mesodermal fates as a whole, a specification of the dwindling mesodermal progenitors into vascular endothelium will not appear as a significant upregulation compared to the wildtype embryos also assayed.

Numerous other genes that relate to formation of mesodermal structures are down regulated, *eve1* and *foxa1*, which work to regulate mesodermal differentiation, and *her1* and *her6* , which assists in somite formation (Henry et al., 2002). Due to the initial allocation of NMPs into neural and mesodermal fates being skewed towards neural ectoderm, it is no surprise that many of the downstream pathways and structures in these cell lineages are also dysregulated.

7.2.7 What is underlying the gene expression changes

I theorise that the shift toward neural processes is due to cohesin dysregulation of *wnt3a*, or genes that interact with *wnt3a*, as opposed to cohesin up-regulation of *sox2*, and the associated neural fate. If *wnt3a* levels are being down-regulated as a result of cohesin loss, either model of neural cell expansion (Figure 7-1 - middle) results in the overall increase (or apparent up-regulation) in *sox2* expression observed in these analyses. The *wnt3a* expression pattern in wildtype and cohesin-deficient embryos, when compared to a schematic outlining the normal tissue distribution in the tailbud (figure 7-1) supports the notion that there is insufficient *wnt3a* present in both NMPs and mesodermal progenitors to commit cells to the correct mesodermal fates in cohesin-deficient embryos, as there is less *wnt3a* expression observed in these regions.

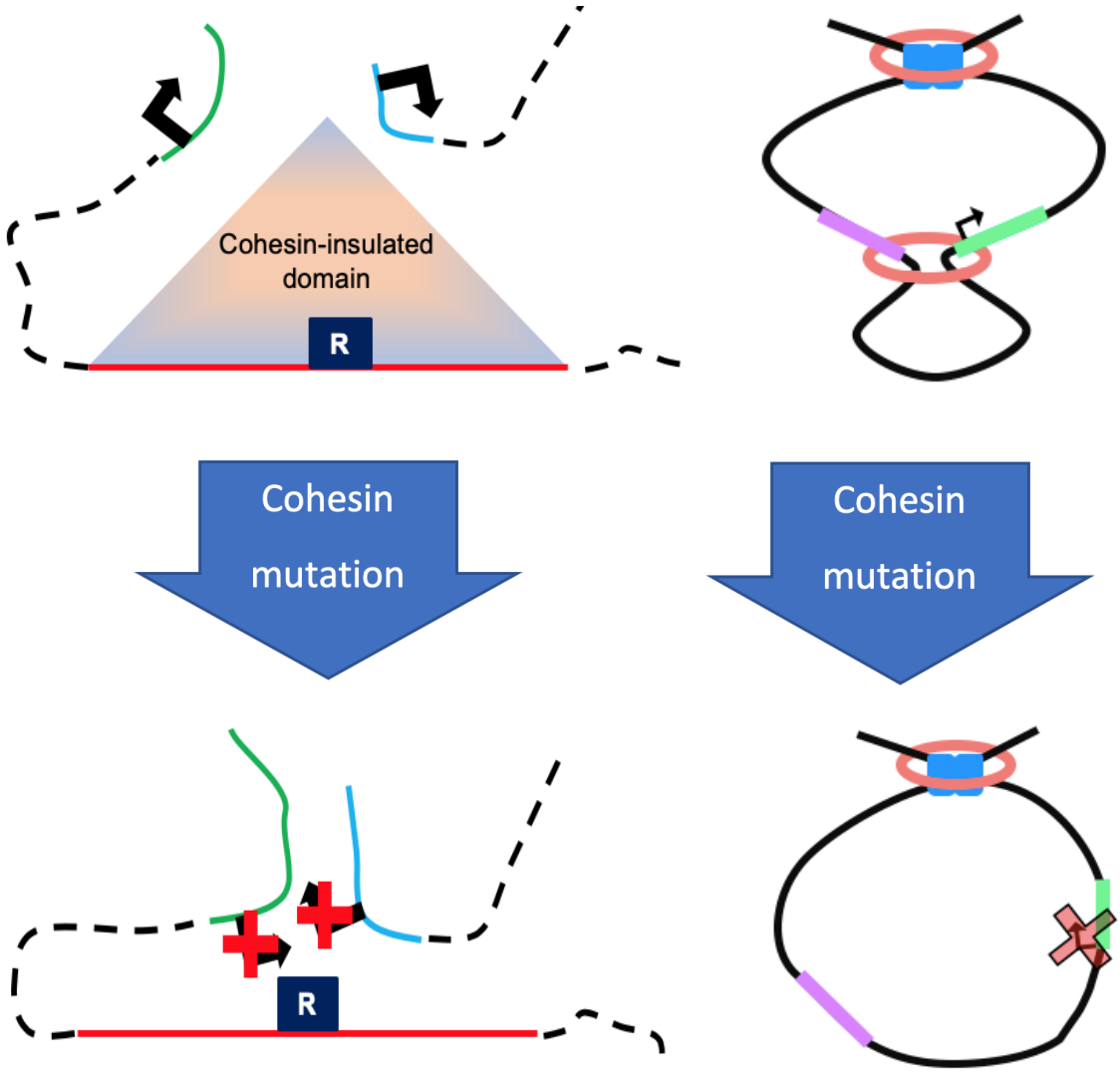


Figure 7-2 - Proposed models of unconstrained chromatin affecting gene expression*. Left panel – Cohesin-deficiency allowing a repressor which was once contained within a cohesin insulated domain to act on multiple genes. Right panel – Cohesin deficiency reducing contacts in the chromatin. Without cohesin regulation the pink enhancer cannot activate expression of the green gene.*

One theory as to how cohesin-deficiency affects genes in development is that cohesin-deficient cells are open to ectopic signalling, through unconstrained chromatin, allowing signals which cannot normally affect expression having an effect. Repressors which are normally sequestered away from *wnt3a* by cohesin organisation of chromatin may now be able to act on *wnt3a* directly, repressing its expression (Figure 7-2 – left panel). Any number of the plethora of gene products that regulate Wnt signalling may also be open to ectopic signals. For example, BMP upregulates Wnt signalling, by repressing Wnt inhibitors

in the very posterior of the tailbud (Row & Kimelman, 2009). We observed a down regulation of multiple BMP genes, which would allow Wnt inhibitor levels to increase, reducing *wnt3a* signal and all downstream pathways.

Another theory is that a lack of cohesin may affect gene expression more directly, through insufficient organisation of local chromatin organisation. The right panel in figure 7-2 depicts a green gene which is expressed when cohesin organisation brings a purple enhancer within its proximity. In a cohesin deficiency, this regulation may not occur, and genes are therefore not expressed. *Wnt3a* expression levels may have dropped in *rad21*^{nz171} homozygote tailbuds due to a lack of cohesin regulating its expression by enhancer interactions.

We aimed to test the unconstrained chromatin theory by comparing expression intensity of the mCherry reporter, and by proxy, Wnt target gene expression in a developing embryo, in transgenic lines of cohesin-deficient vs cohesin-present fish. Unfortunately, this work was not completed in the scope of this masters.

7.2.8 How do our findings relate to the developmental defects seen in humans?

The specific genes that are dysregulated in cohesin-deficient zebrafish embryos do not necessarily directly relate to human cohesinopathy traits. Though both species are vertebrates, zebrafish, which are amniotes, have slightly differing axial extension programmes to mammals, and thus different pathways can be affected by cohesin reduction. Though it is exciting to see which genes are dysregulated in cohesin-deficient zebrafish tailbuds, it does not allow verification of how cohesin regulates these genes. Determining how cohesin regulates the expression of genes to allow these changes in pathways is the next step in deciphering the cohesinopathies.

7.3 Future directions

Many exciting leads came out of these analyses, and examination of proposed ideas could solidify the theories. The ideal next step in examining dysregulated gene expression and cell fate skewing in cohesin-deficient embryos would be to perform double fluorescent *in*

situ hybridisation of *sox2* and *wnt3a* genes in cohesin-deficient compared to wildtype tailbuds. Examining expression of these two genes with Nikon C2 confocal microscopy would allow much higher magnification of the tailbud cells and may allow visualisation of the proposed expansion of neural ectoderm.

Continuing with analysis of the impact of signalling pathways on chromatin is also crucial in elucidating cohesins role in developmental gene expression. To continue this, outcrosses and subsequent BIO-6 stimulations should be performed as outlined in figure 6-1.

7.4 Conclusions

This Masters research first aimed to uncover the functionality of a familial RAD21 variant that may be causative of myeloid dysplasia observed in a cohort of patients. It was determined via a *runx1 in situ* hybridisation assay that the variant does produce a functional RAD21 protein, and protein mapping of the variant site within the 3D RAD21 molecule predicts that the variant does not lie within RAD21 interacting domains. Any effect this variant has on RAD21 functionality in these patients is therefore predicted to be due to changes in variant-cohesin interactions with the genome, which was not further explored in the scope of this Masters.

This research also aimed to explore the effect that cohesin has on regulating gene expression in a bipotential progenitor cell population in the developing embryo. Through a high-throughput RNA-sequencing experiment on the tailbuds of cohesin-deficient zebrafish embryos compared to their wildtype siblings, it was determined that a loss of cohesin resulted in this bipotent stem cell population skewing towards neural ectoderm, at the expense of mesodermal progenitors. This skewing is theorised to be due to a lack of cohesin regulation of *wnt3a* – either directly, or through its numerous regulators. A role for cohesin in regulating gene expression at the translational level was also uncovered, as in cohesin-deficiency, ribosome biogenesis and RNA processes were significantly dysregulated.

The abundance of data generated from the RNA-sequencing analysis opens many new doors for exploration into cohesins role in regulating gene expression throughout development.

APPENDIX 1 – CHEMICALS AND SOLUTIONS

All solutions were made using sterile milli-Q water unless otherwise stated. The pH adjustments were made using 1 M HCl or NaOH.

Luria-Broth (LB) Medium - 4 g Tryptone, 2 g Yeast extract, 4 g Sodium Chloride, make up to 400mL with milli Q water. Sterilise by autoclaving.

LB agar - 4 g Tryptone, 2 g Yeast extract, 4 g Sodium Chloride, 6g bacterial agar, make up to 400mL with milli Q water. Sterilise by autoclaving.

Ethanol - Absolute ethanol dissolved in ddH_2O to final concentrations of 70% and 80%

Glycerol - Glycerol dissolved in UP water to final concentrations of 30% and 80%

Agarose Gels - The desired percentage produced by dissolving agarose powder in 1 x TAE. Per 60 mL gel, 3.6uL of ethidium bromide added.

50 x E3 - 14.6g NaCl, 0.63g KCl, 2.43gCaCl2H2O, 4.07g MgSO47H2O, make up to 1 litre with MilliQ water, and pH to 7.2. Sterilise by autoclaving. E3 used at a 1x concentration.

7%DMSO

100uM BIO-6 in 7% DMSO

List of solutions used in *in situ* hybridisation.

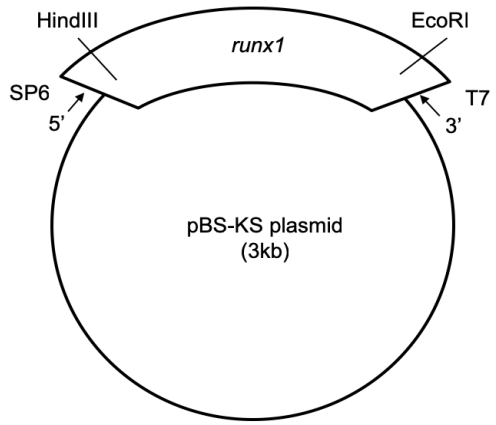
All solutions were autoclaved.

Solution name	Component	Amount
10 x Phosphate buffered saline (PBS) (pH 7.4)	Potassium chloride (KCl)	2.0g
	Potassium phosphate (KH_2PO_4)	2.0g
	Sodium phosphate (NA2PO4-2H2O)	14.4g
	Sodium chloride (NaCl)	80.0g
	MilliQ	1L
1x PBS	10X PBS	100mL
	MilliQ	900mL
1 x Phosphate buffered saline with Tween (PBST)	1 X PBS	1L
	Tween 20	200mL
20x Saline sodium citrate (SSC) buffer (pH 7)	Sodium chloride (NaCl)	175.3g
	Trisodium citrate ($Na_3C_6H_5O_7$)	88.2g
	MilliQ	1L
0.5M Ethylenediaminetetraacetic acid (EDTA)	EDTA	18.6g
	sodium hydroxide (NaOH)	2.0g
	MilliQ	100mL
Hybridisation Buffer (HB+) pH 6.0	formamide	25mL
	20x SSC	12.5mL
	Heparin(5mg/mL)	500mL
	0.5M EDTA	500mL
	tRNA from wheat germ type V (50ug/mL)	250uL
	20% Tween 20	250uL
	1M Citric acid ($C_6H_8O_7$)	460uL
	MilliQ	11.5mL
Hybridisation Buffer (HB-) pH 6.0	formamide	25mL
	20x SSC	12.5mL
	0.5M EDTA	500mL
	20% Tween 20	250uL

	1M Citric acid ($C_6H_8O_7$)	460uL
	MilliQ	11.5mL
10x Maleic acid buffer (MAB)	Maleic acid ($C_4H_4O_4$)	116g
	sodium chloride (NaCl)	87g
	sodium hydroxide (NaOH)	40g
	MilliQ	1L
1 x MAB	10x MAB	100mL
	MilliQ	900mL
2% Roche blocking buffer	Roche blocking agent (2% weight/volume) dissolved at 60 degrees	1g
	1x MAB	50mL
Staining solution	1M Tris pH 9.5	5mL
	2M magnesium chloride (MgCl2)	1.25mL
	5M sodium chloride (NaCl)	1mL
	20% Tween 20	250uL
	MilliQ	42.4uL

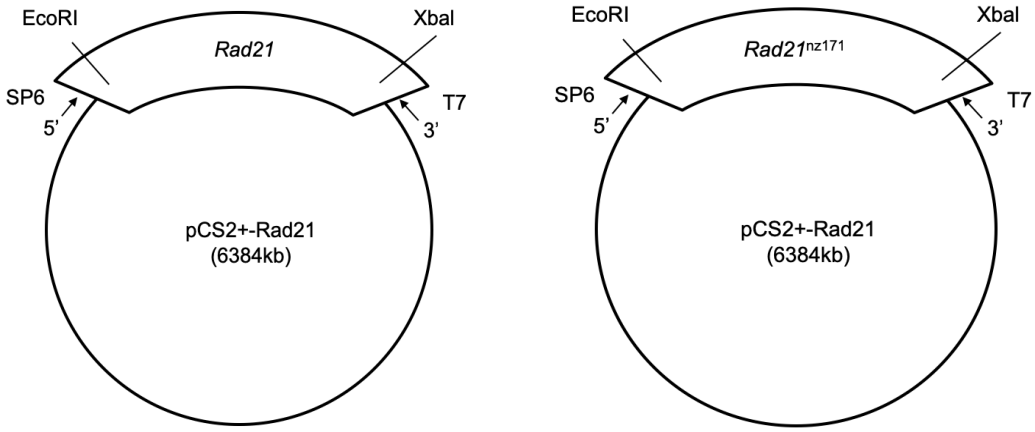
APPENDIX 2 – PLASMID MAPS

Runx1



Depiction of the pBS-KS runx1 plasmid. HINDIII and ECOR1 sites identified. SP6/T3 and T7 primer directions also identified.

WT-*rad21* and *rad21*^{nz171}



Depiction of the PCS2+ rad21 plasmids. Not1, HINDIII and ECOR1 sites identified. SP6 and T7 primer directions also identified. Variant rad21 plasmids created with SDM of PCS2+-Rad2.

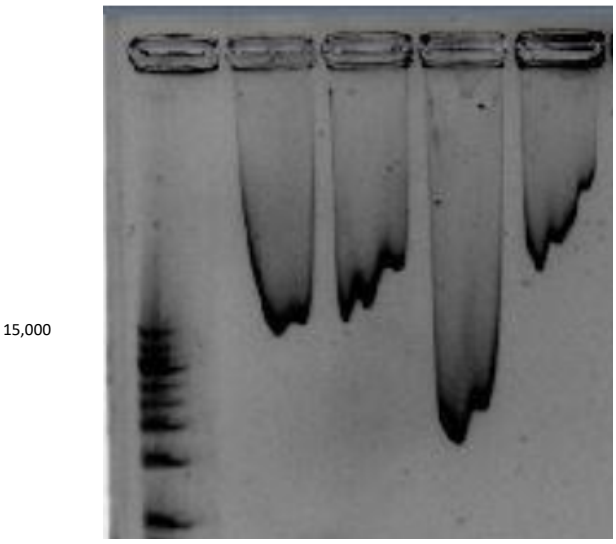
APPENDIX 3 - PRIMER LIST

Details of the primers used in this research are listed below

Primer	**Direction**	**Sequence**
mCherry	Forward	CCCCGTAATGCAGAAGAAGA
mCherry	Reverse	TTGGTCACCTTCAGCTTGG
Stag1b^{nz205}	Forward	GGCGGCTAATAAGAAGGCCA
Stag1b^{nz205}	Reverse	AAGCAGCACACAACCTCGAA
I152F-generating (SDM)	Forward	GTCAGTCATGAGGCTGAAGTTTCCTACTTCCTCCC
I152F-generating (SDM)	Reverse	GGGAGGAAGTAGGAAACTTAGCCTCATGACGAC
S449G-generating (SDM)	Forward	GCCTGCAGCATCCGGGCTCCTCCATG
S449G-generating (SDM)	Reverse	CATGGAGGAGCCCGGTATGCTGCAGC
Detecting I152F mutagenesis	Forward (no reverse required)	GCCTAAGGCGCACAGGTACCAT
Detecting S449G mutagenesis	Forward	GACTGATCCAAGGGCTGAAG
Detecting S449G mutagenesis	Reverse	ACTCGCTGCTTGACTCCACT

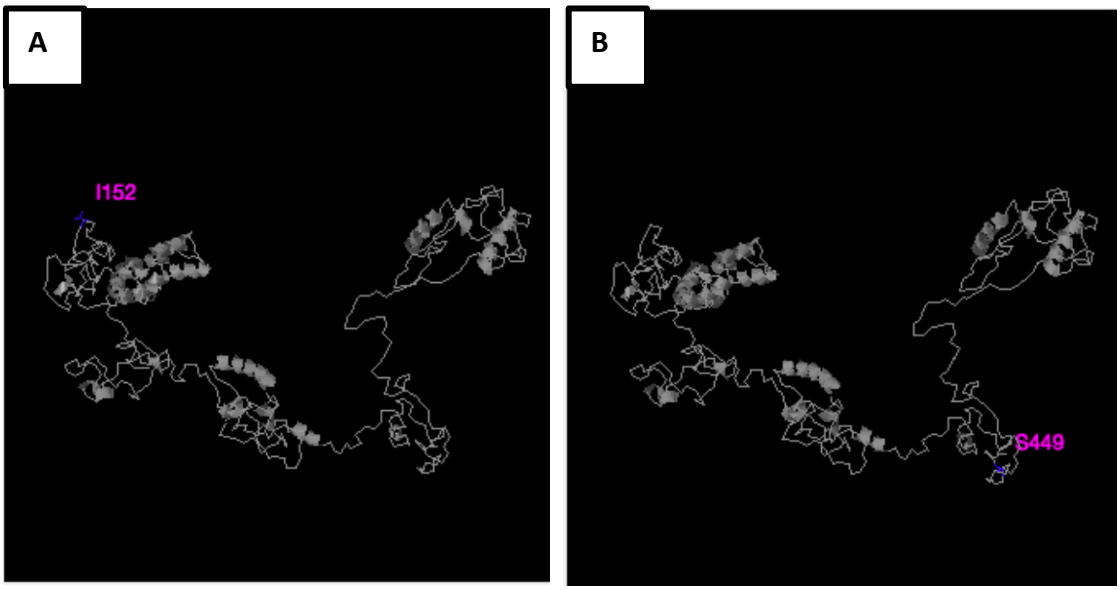
APPENDIX 4 – SUPPLEMENTARY

Control and variant *rad21* mRNA visualisation



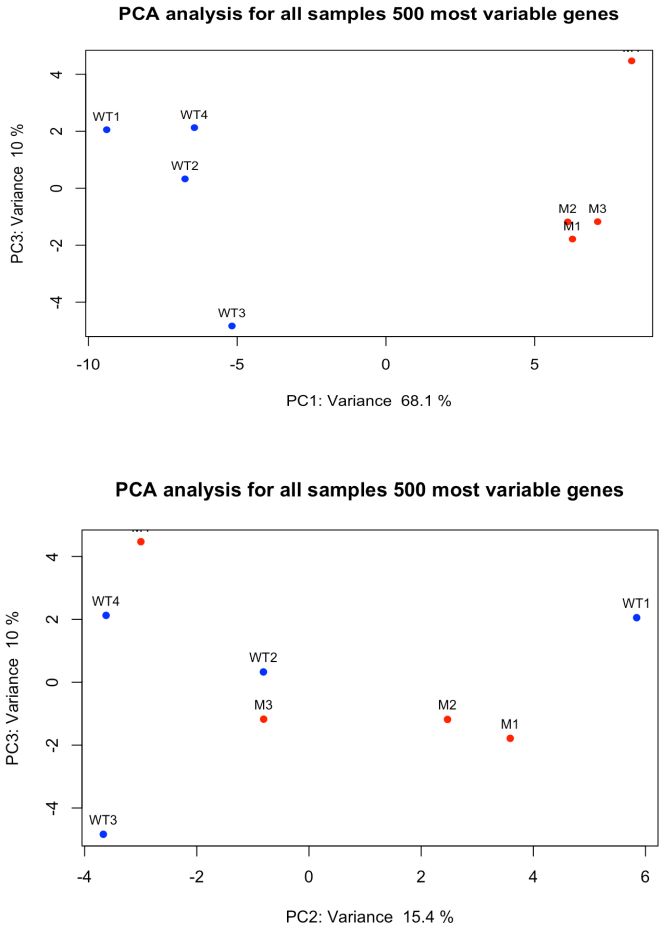
Detecting presence of mRNA on 2% agarose gel run for 2hours at 100V. Lane 1: Invitrogen 1kb ladder, 2: WT-rad21 mRNA, 3: rad21^{nz171} mRNA, 4: I152F-rad21 mRNA, 5: S449G-rad21 mRNA.

STRUM 3D protein prediction



STRUM outputs depicting predicted protein structure for I152F-Rad21 (A) and S449G-Rad21 (B). The variant sites are annotated.

PCA plots demonstrating replicate distinction



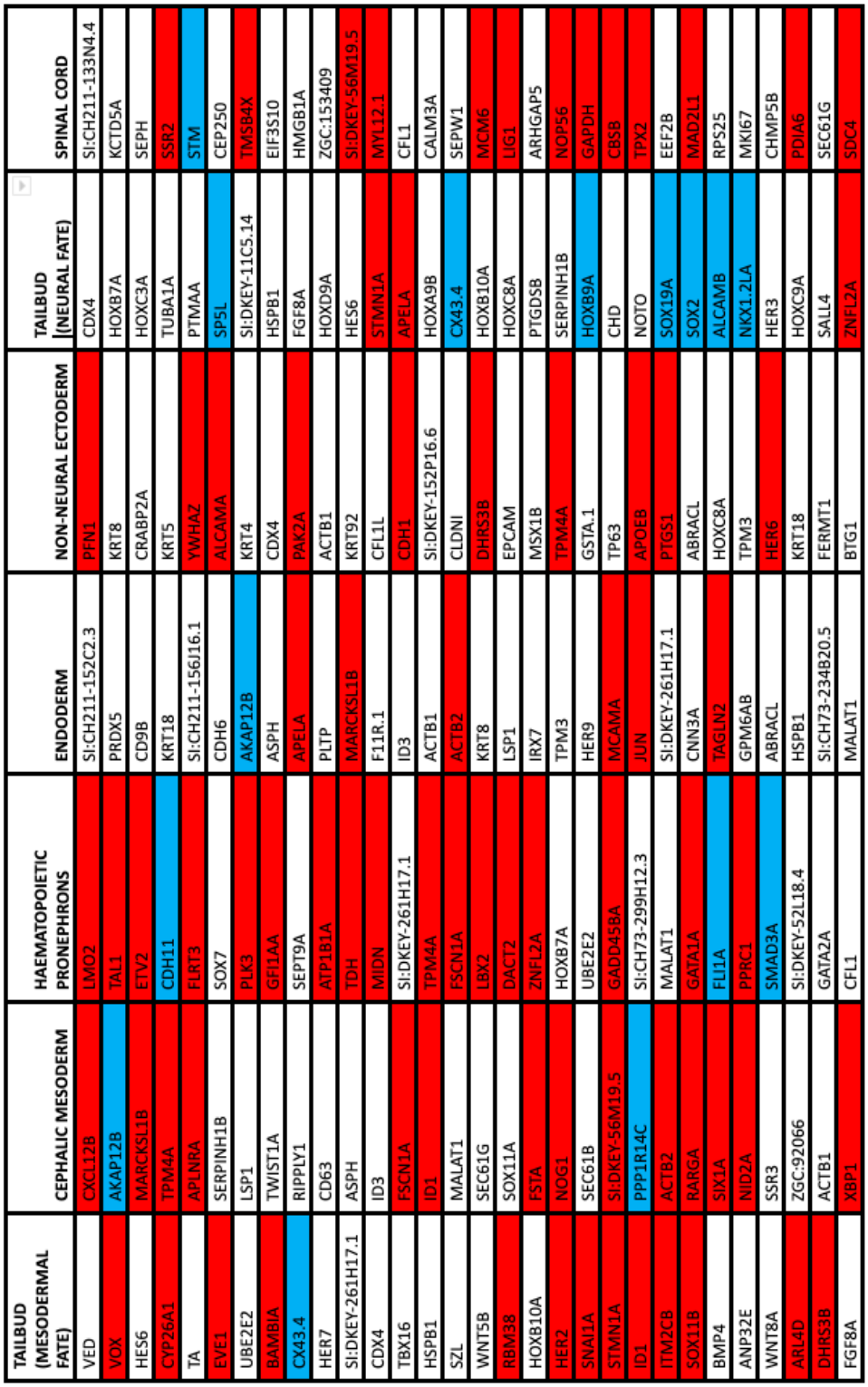
Further PCA plots, demonstrating similarity between replicates. The four wildtype replicates are depicted in blue, and the four rad21^{nz171} homozygote replicates, in red.

Gene dysregulation across zebrafish chromosomes

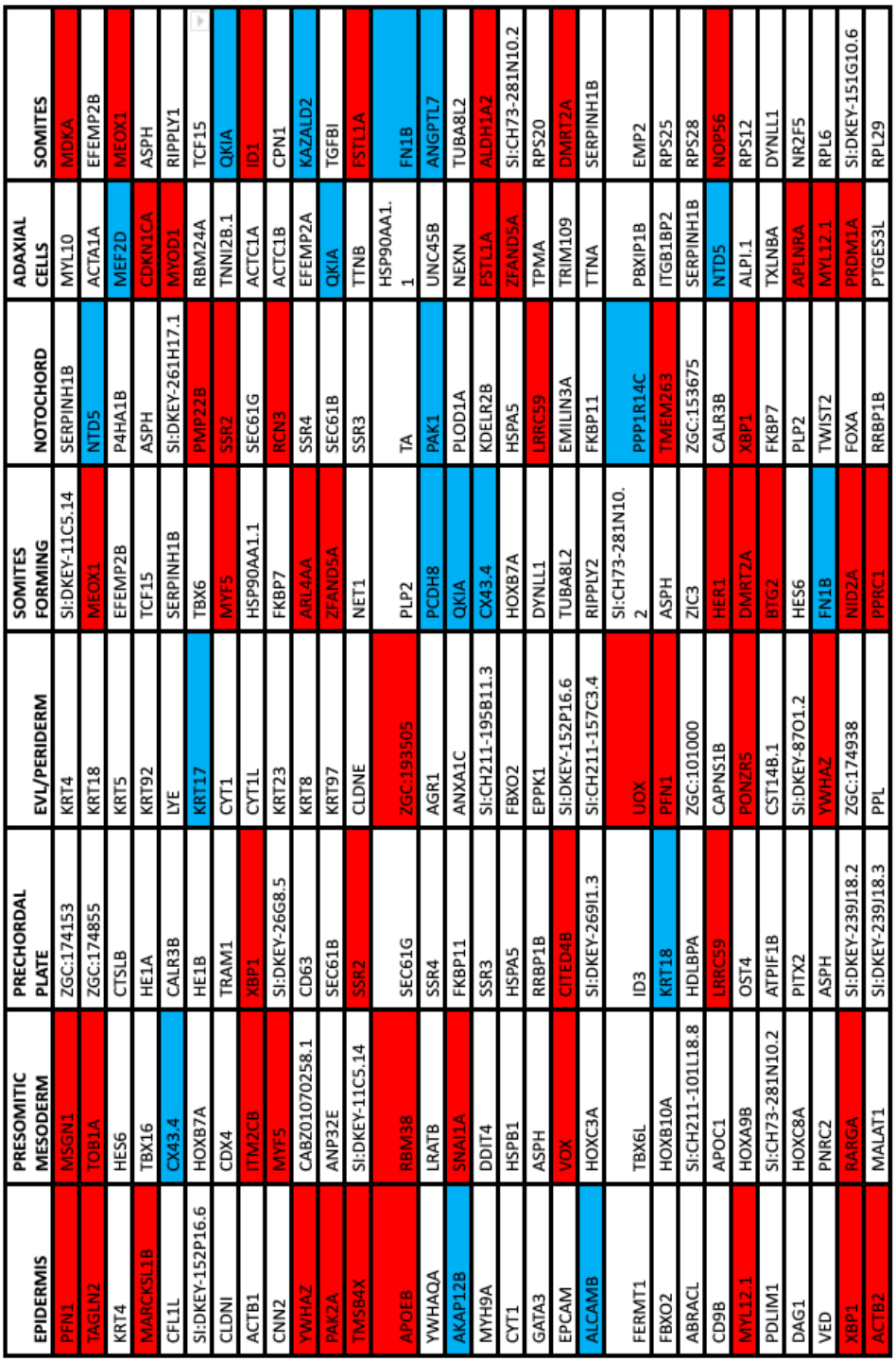
	chromosome	observed	expected	odds.ratio	pvalue	padj
1	1	35	30.2	1.17174182178259	0.206560317370163	0.730503803266424
2	2	26	33.6	0.760119947104394	0.931483565982069	0.999999968488283
3	3	32	34.1	0.932692860398324	0.673468864103153	0.957167318883934
4	4	18	49.4	0.341935981410058	0.999999968488283	0.999999968488283
5	5	32	34.3	0.927843277077232	0.683888807919596	0.957167318883934
6	6	28	26.8	1.04600342375757	0.436625960530034	0.901813913982312
7	7	38	36.1	1.05856077627141	0.392300285210108	0.901813913982312
8	8	25	26.5	0.939589978267278	0.647188966145488	0.957167318883934
9	9	30	23.7	1.28765420466681	0.111212883932436	0.722883745560834
10	10	28	24	1.17607883325397	0.229417873165557	0.730503803266424
11	11	20	23.7	0.835004100368652	0.81281620719439	0.966319336351284
12	12	33	21.1	1.61522787339191	0.00805489880287973	0.104713684437436
13	13	21	23	0.906523487767016	0.69946842533826	0.957167318883934
14	14	24	21	1.15098833898167	0.280963001256317	0.730503803266424
15	15	21	24.9	0.834894638505962	0.817654823066471	0.966319336351284
16	16	41	26.7	1.58831285099624	0.00475330110639787	0.104713684437436
17	17	22	23.2	0.943988244481981	0.633863284544477	0.957167318883934
18	18	22	21.1	1.04526484868087	0.450906956991156	0.901813913982312
19	19	36	24.2	1.53129522549294	0.0123654705988276	0.107167411856505
20	20	31	26.5	1.18397965798401	0.206520603110775	0.730503803266424
21	21	27	23.5	1.15885026150206	0.256292432758985	0.730503803266424
22	22	19	25.5	0.734904651311146	0.927462405907933	0.999999968488283
23	23	15	22.9	0.643373096759998	0.970414589755677	0.999999968488283
24	24	22	17.6	1.26452202822571	0.170003060515323	0.730503803266424
25	25	18	20.5	0.871612789717486	0.747060051115944	0.966319336351284
26	MT	1	0.8	1.31173204016588	0.539780799812089	0.957167318883934

Table listing all zebrafish chromosomes and the expected and observed amounts of gene dysregulation in the RNAseq analyses. Chromosome 12 and 16 both have a statistically significantly increased amount of gene dysregulation.

Comparing gene lists to other research



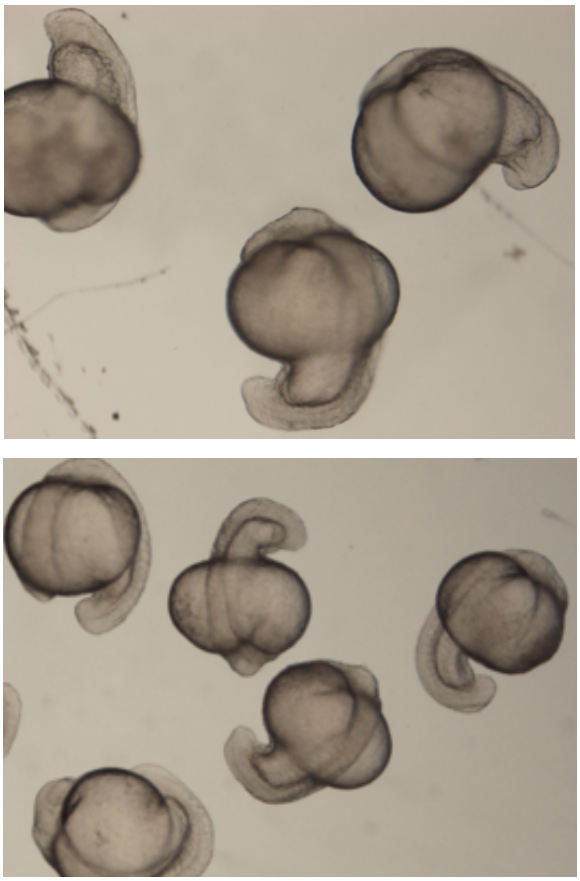
TAILBUD (MESODERMAL FATE)	CEPHALIC MESODERM	HAEMATOPOIETIC PRONEPHRONS	ENDODERM	NON-NEURAL ECTODERM	TAILBUD (NEURAL FATE)	SPINAL CORD
VED	CXCL12B	LMO2	SI:CH211-152C2.3	PFN1	CDX4	SI:CH211-133N4.4
VOX	AKAP12B	TAL1	PRDX5	KRT8	HOXB7A	KCTD5A
HES6	MARCKSL1B	ETV2	CD9B	CRABP2A	HOXC3A	SEPH
CYP26A1	TPM4A	CDH11	KRT18	KRT5	TUBA1A	SSR2
TA	APLNRA	FLRT3	SI:CH211-156J16.1	YWHAZ	PTMAA	STM
EVE1	SERPINH1B	SOX7	CDH6	ALCAMA	SP5L	CEP250
UBE2E2	LSP1	PLK3	AKAP12B	KRT4	SI:DKEY-11C5.14	TMSB4X
BAMBIA	TWIST1A	GFI1AA	ASPH	CDX4	HSPB1	EIF3S10
CX43.4	RIPPLY1	SEPT9A	APELA	PAK2A	FGF8A	HMGB1A
HER7	CD63	ATP1B1A	PLTP	ACTB1	HOXD9A	ZGC:153409
SI:DKEY-261H17.1	ASPH	TDH	MARCKSL1B	KRT92	HES6	SI:DKEY-56M19.5
CDX4	ID3	MIDN	F11R.1	CFL1L	STMN1A	MYL12.1
TBX16	FSCN1A	SI:DKEY-261H17.1	ID3	CDH1	APELA	CFL1
HSPB1	ID1	TPM4A	ACTB1	SI:DKEY-152P16.6	HOXA9B	CALM3A
SZL	MALAT1	FSCN1A	ACTB2	CLDNI	CX43.4	SEPW1
WNT5B	SEC61G	LBX2	KRT8	DHRS3B	HOXB10A	MCM6
RBM38	SOX11A	DACT2	LSP1	EPCAM	HOXC8A	LIG1
HOXB10A	FSTA	ZNFL2A	IRX7	MSX1B	PTGDSB	ARHGAP5
HER2	NOG1	HOXB7A	TPM3	TPM4A	SERPINH1B	NOP56
SNAI1A	SEC61B	UBE2E2	HER9	GSTA.1	HOXB9A	GAPDH
STMN1A	SI:DKEY-56M19.5	GADD45BA	MCAMA	TP63	CHD	CBSB
ID1	PPP1R14C	SI:CH73-299H12.3	JUN	APOEB	NOTO	TPX2
ITM2CB	ACTB2	MALAT1	SI:DKEY-261H17.1	PTGS1	SOX19A	EEF2B
SOX11B	RARGA	GATA1A	CNN3A	ABRACL	SOX2	MAD2L1
BMP4	SIX1A	FLI1A	TAGLN2	HOXC8A	ALCAMB	RPS25
ANP32E	NID2A	PPRC1	GPM6AB	TPM3	NKX1.2LA	MKI67
WNT8A	SSR3	SMAD3A	ABRACL	HER6	HER3	CHMP5B
ARL4D	ZGC:92066	SI:DKEY-52L18.4	HSPB1	KRT18	HOXC9A	PDIA6
DHRS3B	ACTB1	GATA2A	SI:CH73-234B20.5	FERMT1	SALL4	SEC61G
FGF8A	XBP1	CFL1	MALAT1	BTG1	ZNFL2A	SDC4



EPIDERMIS	PRESOMITIC MESODERM	PRECHORDAL PLATE	EVL/PERIDERM	SOMITES FORMING	NOTOCHORD	ADAXIAL CELLS	SOMITES
PFN1	MSGN1	ZGC:174153	KRT4	SI:DKEY-11C5.14	SERPINH1B	MY110	MDKA
TAGLN2	TOB1A	ZGC:174855	KRT18	MEOX1	NTD5	ACTA1A	EFEMP2B
KRT4	HES6	CTSLB	KRT5	EFEMP2B	P4HA1B	MEF2D	MEOX1
MARCKSL1B	TBX16	HE1A	KRT92	TCF15	ASPH	CDKN1CA	ASPH
CFL1L	CX43.4	CALR3B	LYE	SERPINH1B	SI:DKEY-261H17.1	MYOD1	RIPPLY1
SI:DKEY-152P16.6	HOXB7A	HE1B	KRT17	TBX6	PMP22B	RBM24A	TCF15
CLDNI	CDX4	TRAM1	CYT1	MYF5	SSR2	TNNI2B.1	QKIA
ACTB1	ITM2CB	XBP1	CYT1L	HSP90AA1.1	SEC61G	ACTC1A	ID1
CNN2	MYF5	SI:DKEY-26G8.5	KRT23	FKBP7	RCN3	ACTC1B	CPN1
YWHAZ	CABZ01070258.1	CD63	KRT8	ARL4AA	SSR4	EFEMP2A	KAZALD2
PAK2A	ANP32E	SEC61B	KRT97	ZFAND5A	SEC61B	QKIA	TGFBI
TMSB4X	SI:DKEY-11C5.14	SSR2	CLDNE	NET1	SSR3	TTNB	FSTL1A
APOEB	RBM38	SEC61G	ZGC:193505	PLP2	TA	HSP90AA1.1	FN1B
YWHAQA	LRATB	SSR4	AGR1	PCDH8	PAK1	UNC45B	ANGPTL7
AKAP12B	SNAI1A	FKBP11	ANXA1C	QKIA	PLOD1A	NEXN	TUBA8L2
MYH9A	DDIT4	SSR3	SI:CH211-195B11.3	CX43.4	KDELR2B	FSTL1A	ALDH1A2
CYT1	HSPB1	HSPA5	FBXO2	HOXB7A	HSPA5	ZFAND5A	SI:CH73-281N10.2
GATA3	ASPH	RRBP1B	EPPK1	DYNLL1	LRRC59	TPMA	RPS20
EPCAM	VOX	CITED4B	SI:DKEY-152P16.6	TUBA8L2	EMILIN3A	TRIM109	DMRT2A
ALCAMB	HOXC3A	SI:DKEY-269I1.3	SI:CH211-157C3.4	RIPPLY2	FKBP11	TTNA	SERPINH1B
FERMT1	TBX6L	ID3	UOX	SI:CH73-281N10.2	PPP1R14C	PBXIP1B	EMP2
FBXO2	HOXB10A	KRT18	PFN1	ASPH	TMEM263	ITGB1BP2	RPS25
ABRACL	SI:CH211-101L18.8	HDLBPA	ZGC:101000	ZIC3	ZGC:153675	SERPINH1B	RPS28
CD9B	APOC1	LRRC59	CAPNS1B	HER1	CALR3B	NTD5	NOP56
MYL12.1	HOXA9B	OST4	PON2R5	DMRT2A	XBP1	ALPI.1	RPS12
PDLIM1	SI:CH73-281N10.2	ATPIF1B	CST14B.1	BTG2	FKBP7	TXLNBA	DYNLL1
DAG1	HOXC8A	PITX2	SI:DKEY-87O1.2	HES6	PLP2	APLNRA	NR2F5
VED	PNRC2	ASPH	YWHAZ	FN1B	TWIST2	MYL12.1	RPL6
XBP1	RARGA	SI:DKEY-239J18.2	ZGC:174938	NID2A	FOXA	PRDM1A	SI:DKEY-151G10.6
ACTB2	MALAT1	SI:DKEY-239J18.3	PPL	PPRC1	RRBP1B	PTGES3L	RPL29

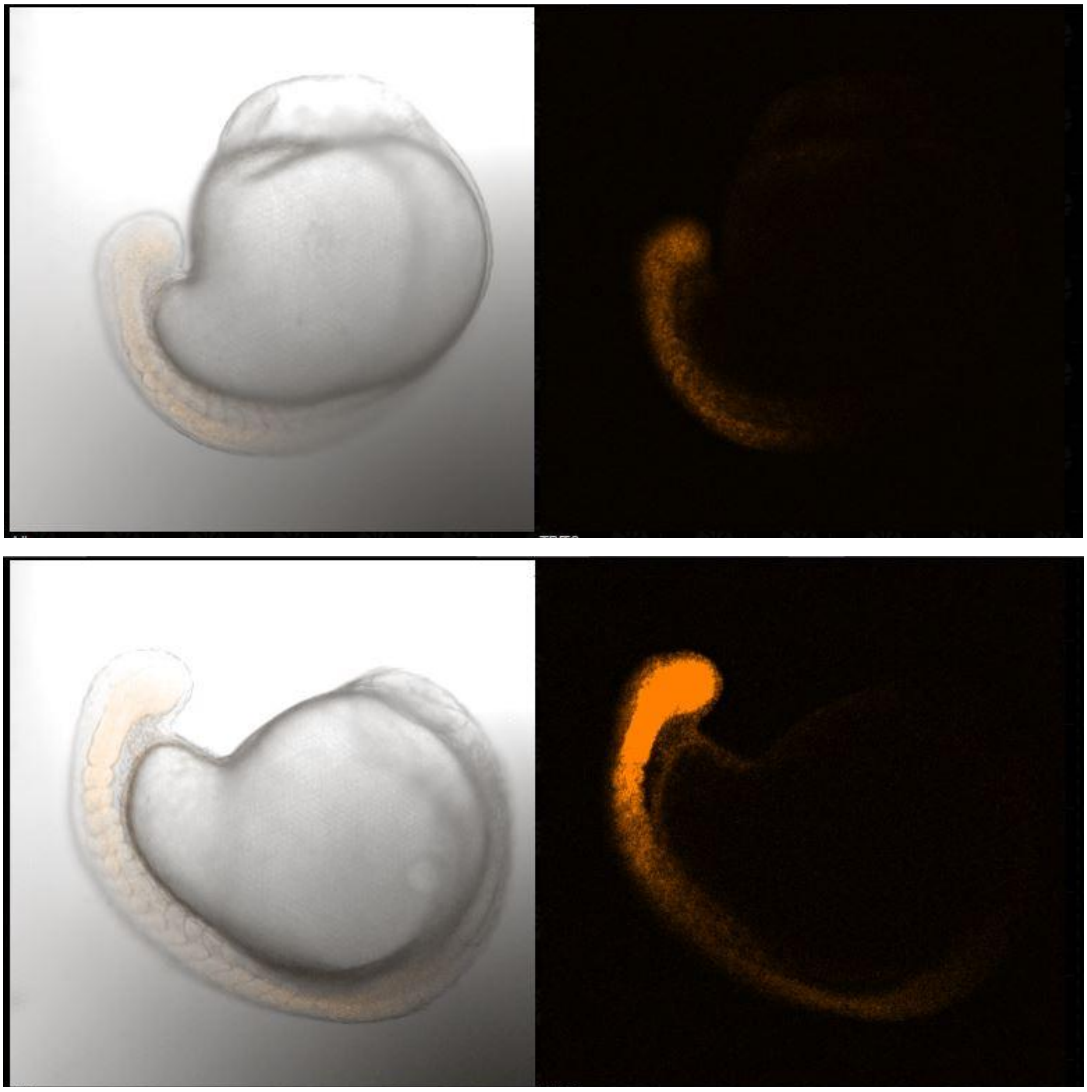
Table adapted from Farrell et al supplementary. These gene lists were created to demonstrate the key genes underlying distinct cell trajectories. Red highlighted genes were downregulated in our RNAseq analyses. Blue highlighted genes were upregulated in our RNAseq analyses

Testing BIO-6 toxicity



Embryos imaged after stimulation with BIO-6 for 12 hours. Top. Embryos after stimulation with 5 uM BIO-6. The tailbuds of the embryos are stumpy in comparison to those treated with 2.5 uM BIO-6, and the embryos are dying. Bottom - Embryos after stimulation with 2.5 uM BIO-6. Development of the embryo appears normal.

Examining mCherry fluorescence



MCherry fluorescence in embryos from a Tg(7xTCFXla.Siam:nlsmCherry)ia5 heterozygote/ rad21^{nz171} heterozygote fish in-cross at 18 hpf. There is clear distinction between offspring embryos carrying one copy of the transgene (top) and two copies of the transgene (bottom) based upon the degree of fluorescence.

Rad21 mutants in mitotic delay

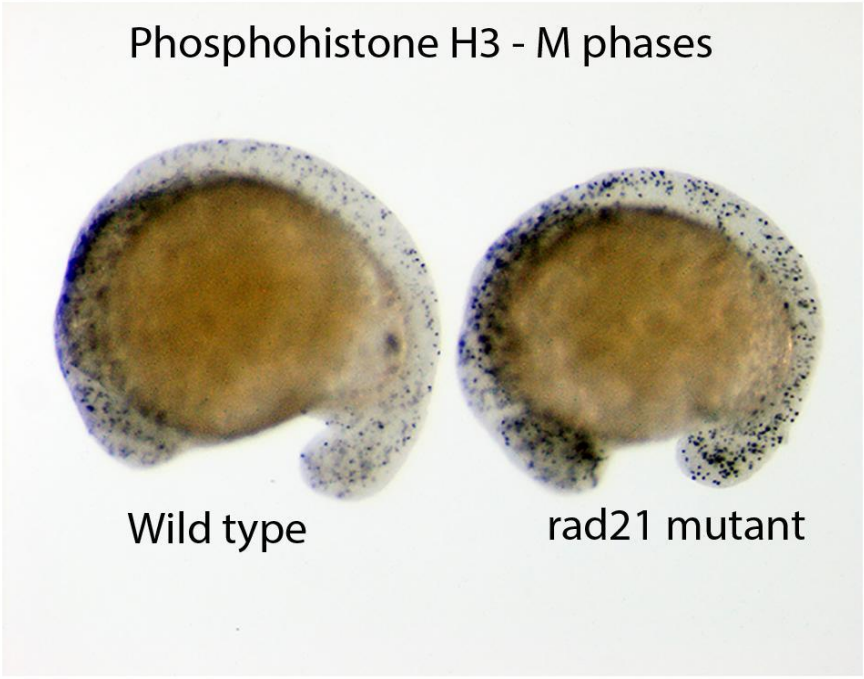


Image adapted from Horsfield et al 2007. Phosphohistine H3 labelling of M phase shows more cells with mitotic delay in rad21 mutants compared with wildtype.

REFERENCES

1. Ali, T., Renkawitz, R., & Bartkuhn, M. (2016). Insulators and domains of gene expression. *Current Opinion in Genetics & Development*, *37*, 17–26. https://doi.org/10.1016/j.gde.2015.11.009

2. Attardi, A., Fulton, T., Florescu, M., Shah, G., Muresan, L., Lenz, M. O., Lancaster, C., Huisken, J., Oudenaarden, A. van, & Steventon, B. (2018). Neuromesodermal progenitors are a conserved source of spinal cord with divergent growth dynamics. *Development*, *145*(21). https://doi.org/10.1242/dev.166728

3. Banerji, R., Skibbens, R. V., & Iovine, M. K. (2017). Cohesin mediates Esco2-dependent transcriptional regulation in a zebrafish regenerating fin model of Roberts Syndrome. *Biology Open*, *6*(12), 1802–1813. https://doi.org/10.1242/bio.026013

4. Baquero-Montoya, C., Gil-Rodríguez, M. C., Teresa-Rodrigo, M. E., Hernández-Marcos, M., Bueno-Lozano, G., Bueno-Martínez, I., Remeseiro, S., Fernández-Hernández, R., Bassecourt-Serra, M., Alba, M. R. de, Queralt, E., Losada, A., Puisac, B., Ramos, F. J., & Pié, J. (2014). Could a patient with SMC1A duplication be classified as a human cohesinopathy? *Clinical Genetics*, *85*(5), 446–451. https://doi.org/10.1111/cge.12194

5. Barrington, C., Finn, R., & Hadjur, S. (2017). Cohesin biology meets the loop extrusion model. *Chromosome Research*, *25*(1), 51–60. https://doi.org/10.1007/s10577-017-9550-3

6. Barrington, C., Georgopoulou, D., Pezic, D., Varsally, W., Herrero, J., & Hadjur, S. (2019). Enhancer accessibility and CTCF occupancy underlie asymmetric TAD architecture and cell type specific genome topology. *Nature Communications*, *10*(1), 1–14. https://doi.org/10.1038/s41467-019-10725-9

7. Basel-Vanagaite, L., Wolf, L., Orin, M., Larizza, L., Gervasini, C., Krantz, I. D., & Deardoff, M. A. (2016). Recognition of the Cornelia de Lange syndrome phenotype with facial dysmorphology novel analysis. *Clinical Genetics*, *89*(5), 557–563. https://doi.org/10.1111/cge.12716

8. Bergeron, M., Chang, K., & Ishman, S. L. (2019). Cornelia de lange manifestations in otolaryngology: A systematic review and meta-analysis. *The Laryngoscope*, *n/a*(n/a). https://doi.org/10.1002/lary.28169

9. Birkenbihl, R. P., & Subramani, S. (1992). Cloning and characterization of rad21 an essential gene of Schizosaccharomyces pombe involved in DNA double-strand-break repair. *Nucleic Acids Research*, *20*(24), 6605–6611. https://doi.org/10.1093/nar/20.24.6605

10. Bose, T., Lee, K. K., Lu, S., Xu, B., Harris, B., Slaughter, B., Unruh, J., Garrett, A., McDowell, W., Box, A., Li, H., Peak, A., Ramachandran, S., Seidel, C., & Gerton, J. L. (2012). Cohesin Proteins Promote Ribosomal RNA Production and Protein Translation in Yeast and Human Cells. *PLoS Genetics*, *8*(6). https://doi.org/10.1371/journal.pgen.1002749

11. Boyle, M. I., Jespersgaard, C., Nazaryan, L., Bisgaard, A.-M., & Tümer, Z. (2017). A novel RAD21 variant associated with intrafamilial phenotypic variation in Cornelia de Lange syndrome – review of the literature. *Clinical Genetics*, *91*(4), 647–649. https://doi.org/10.1111/cge.12863

12. Butler, M. T., & Wallingford, J. B. (2017). Planar cell polarity in development and disease. *Nature Reviews Molecular Cell Biology*, *18*(6), 375–388. https://doi.org/10.1038/nrm.2017.11

13. Casa, V., Gines, M. M., Gusmao, E. G., Slotman, J. A., Zirkel, A., Josipovic, N., Oole, E., IJcken, W. F. J. van, Houtsmuller, A. B., Papantonis, A., & Wendt, K. S. (2019). Redundant and specific roles of cohesin STAG subunits in chromatin looping and transcription control. *BioRxiv*, 642959. https://doi.org/10.1101/642959

14. Chavda, A. P., Ang, K., & Ivanov, D. (2017). The torments of the cohesin ring. *Nucleus*, *8*(3), 261–267. https://doi.org/10.1080/19491034.2017.1295200

15. Chen, X., Ke, Y., Wu, K., Zhao, H., Sun, Y., Gao, L., Liu, Z., Zhang, J., Tao, W., Hou, Z., Liu, H., Liu, J., & Chen, Z.-J. (2019). Key role for CTCF in establishing chromatin structure in human embryos. *Nature*, *576*(7786), 306–310. https://doi.org/10.1038/s41586-019-1812-0

16. Cremer, T., Cremer, M., Hübner, B., Strickfaden, H., Smeets, D., Popken, J., Sterr, M., Markaki, Y., Rippe, K., & Cremer, C. (2015). The 4D nucleome: Evidence for a dynamic nuclear landscape based on co-aligned active and inactive nuclear compartments. *FEBS Letters*, *589*(20, Part A), 2931–2943. https://doi.org/10.1016/j.febslet.2015.05.037

17. Cucco, F., & Musio, A. (2016). Genome stability: What we have learned from cohesinopathies. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *172*(2), 171–178. https://doi.org/10.1002/ajmg.c.31492

18. Darras, S., Fritzenwanker, J. H., Uhlinger, K. R., Farrelly, E., Pani, A. M., Hurley, I. A., Norris, R. P., Osovitz, M., Terasaki, M., Wu, M., Aronowicz, J., Kirschner, M., Gerhart, J. C., & Lowe, C. J. (2018). Anteroposterior axis patterning by early canonical Wnt signaling during hemichordate development. *PLOS Biology*, *16*(1), e2003698. https://doi.org/10.1371/journal.pbio.2003698

19. Davidson, I. F., Bauer, B., Goetz, D., Tang, W., Wutz, G., & Peters, J.-M. (2019). DNA loop extrusion by human cohesin. *Science*, *366*(6471), 1338–1345. https://doi.org/10.1126/science.aaz3418

20. De Kouchkovsky, I., & Abdul-Hay, M. (2016). 'Acute myeloid leukemia: A comprehensive review and 2016 update.' *Blood Cancer Journal*, *6*(7), e441–e441. https://doi.org/10.1038/bcj.2016.50

21. Deardorff, M. A., Wilde, J. J., Albrecht, M., Dickinson, E., Tennstedt, S., Braunholz, D., Mönnich, M., Yan, Y., Xu, W., Gil-Rodríguez, M. C., Clark, D., Hakonarson, H., Halbach, S., Michelis, L. D., Rampuria, A., Rossier, E., Spranger, S., Van Maldergem, L., Lynch, S. A., … Kaiser, F. J. (2012). RAD21 Mutations Cause a Human Cohesinopathy. *The American Journal of Human Genetics*, *90*(6), 1014–1027. https://doi.org/10.1016/j.ajhg.2012.04.019

22. Dixon, J. R., Gorkin, D. U., & Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. *Molecular Cell*, *62*(5), 668–680. https://doi.org/10.1016/j.molcel.2016.05.018

23. Dooley, K., & Zon, L. I. (2000). Zebrafish: A model system for the study of human disease. *Current Opinion in Genetics & Development*, *10*(3), 252–256. https://doi.org/10.1016/S0959-437X(00)00074-5

24. Dorsett, D. (2019). The Many Roles of Cohesin in Drosophila Gene Transcription. *Trends in Genetics*, *35*(7), 542–551. https://doi.org/10.1016/j.tig.2019.04.002

25. Dorsett, D., & Merkenschlager, M. (2013). Cohesin at active genes: A unifying theme for cohesin and gene expression from model organisms to humans. *Current Opinion in Cell Biology*, *25*(3), 327–333. https://doi.org/10.1016/j.ceb.2013.02.003

26. Dorsett, D., & Ström, L. (2012). The Ancient and Evolving Roles of Cohesin in Gene Expression and DNA Repair. *Current Biology*, *22*(7), R240–R250. https://doi.org/10.1016/j.cub.2012.02.046

27. Dorval, S., Masciadri, M., Mathot, M., Russo, S., Revencu, N., & Larizza, L. (2020). A novel RAD21 mutation in a boy with mild Cornelia de Lange presentation: Further delineation of the phenotype. *European Journal of Medical Genetics*, *63*(1), 103620. https://doi.org/10.1016/j.ejmg.2019.01.010

28. Dowen, J. M., Fan, Z. P., Hnisz, D., Ren, G., Abraham, B. J., Zhang, L. N., Weintraub, A. S., Schuijers, J., Lee, T. I., Zhao, K., & Young, R. A. (2014). Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes. *Cell*, *159*(2), 374–387. https://doi.org/10.1016/j.cell.2014.09.030

29. Dunty, W. C., Biris, K. K., Chalamalasetty, R. B., Taketo, M. M., Lewandoski, M., & Yamaguchi, T. P. (2008). Wnt3a/β-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. *Development*, *135*(1), 85–94. https://doi.org/10.1242/dev.009266

30. Eng, T., Guacci, V., & Koshland, D. (2015). Interallelic complementation provides functional evidence for cohesin–cohesin interactions on DNA. *Molecular Biology of the Cell*, *26*(23), 4224–4235. https://doi.org/10.1091/mbc.e15-06-0331

31. Farrar, J. E., Schuback, H. L., Ries, R. E., Wai, D., Hampton, O. A., Trevino, L. R., Alonzo, T. A., Auvil, J. M. G., Davidsen, T. M., Gesuwan, P., Hermida, L., Muzny, D. M., Dewal, N., Rustagi, N., Lewis, L. R., Gamis, A. S., Wheeler, D. A., Smith, M. A., Gerhard, D. S., & Meshinchi, S. (2016). Genomic Profiling of Pediatric Acute Myeloid Leukemia Reveals a Changing Mutational Landscape from Disease Diagnosis to Relapse. *Cancer Research*, *76*(8), 2197–2205. https://doi.org/10.1158/0008-5472.CAN-15-1015

32. Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science*, *360*(6392). https://doi.org/10.1126/science.aar3131

33. Fatica, A., & Tollervey, D. (2002). Making ribosomes. *Current Opinion in Cell Biology*, *14*(3), 313–318. https://doi.org/10.1016/S0955-0674(02)00336-8

34. Fisher, J. B., McNulty, M., Burke, M. J., Crispino, J. D., & Rao, S. (2017). Cohesin Mutations in Myeloid Malignancies. *Trends in Cancer*, *3*(4), 282–293. https://doi.org/10.1016/j.trecan.2017.02.006

35. Freed, E. F., Bleichert, F., Dutca, L. M., & Baserga, S. J. (2010). When ribosomes go bad: Diseases of ribosome biogenesis. *Molecular BioSystems*, *6*(3), 481–493. https://doi.org/10.1039/b919670f

36. Fritz, A. J., Barutcu, A. R., Martin-Buley, L., Wijnen, A. J. van, Zaidi, S. K., Imbalzano, A. N., Lian, J. B., Stein, J. L., & Stein, G. S. (2016). Chromosomes at Work: Organization of Chromosome Territories in the Interphase Nucleus. *Journal of Cellular Biochemistry*, *117*(1), 9–19. https://doi.org/10.1002/jcb.25280

37. Fritz, A. J., Sehgal, N., Pliss, A., Xu, J., & Berezney, R. (2019). Chromosome territories and the global regulation of the genome. *Genes, Chromosomes and Cancer*, *58*(7), 407–426. https://doi.org/10.1002/gcc.22732

38. Fröhling, S., Scholl, C., Gilliland, D. G., & Levine, R. L. (2005). Genetics of Myeloid Malignancies: Pathogenetic and Clinical Implications. *Journal of Clinical Oncology*, *23*(26), 6285–6295. https://doi.org/10.1200/JCO.2005.05.010

39. Gangat, N., Patnaik, M. M., & Tefferi, A. (2016). Myelodysplastic syndromes: Contemporary review and how we treat. *American Journal of Hematology*, *91*(1), 76–89. https://doi.org/10.1002/ajh.24253

40. Gligoris, T. G. (2018). Chromosome Biology: The Smc–Kleisin Enzymology Finally Comes of Age. *Current Biology*, *28*(10), R612–R614. https://doi.org/10.1016/j.cub.2018.03.056

41. Gligoris, T., & Löwe, J. (2016). Structural Insights into Ring Formation of Cohesin and Related Smc Complexes. *Trends in Cell Biology*, *26*(9), 680–693. https://doi.org/10.1016/j.tcb.2016.04.002

42. Goto, H., Kimmey, S. C., Row, R. H., Matus, D. Q., & Martin, B. L. (2017). FGF and canonical Wnt signaling cooperate to induce paraxial mesoderm from tailbud neuromesodermal progenitors through regulation of a two-step epithelial to mesenchymal transition. *Development*, *144*(8), 1412–1424. https://doi.org/10.1242/dev.143578

43. Gouti, M., Metzis, V., & Briscoe, J. (2015). The route to spinal cord cell types: A tale of signals and switches. *Trends in Genetics*, *31*(6), 282–289. https://doi.org/10.1016/j.tig.2015.03.001

44. Gouti, M., Tsakiridis, A., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V., & Briscoe, J. (2014). In Vitro Generation of Neuromesodermal Progenitors Reveals Distinct

Roles for Wnt Signalling in the Specification of Spinal Cord and Paraxial Mesoderm Identity. *PLoS Biology*, *12*(8), e1001937. https://doi.org/10.1371/journal.pbio.1001937

45. Haering, C. H., Schoffnegger, D., Nishino, T., Helmhart, W., Nasmyth, K., & Löwe, J. (2004). Structure and stability of cohesin's smc1-kleisin interaction. *Molecular Cell*, *15*(6), 951–964. https://doi.org/10.1016/j.molcel.2004.08.030

46. Hagstrom, K. A., & Meyer, B. J. (2003). Condensin and cohesin: More than chromosome compactor and glue. *Nature Reviews Genetics*, *4*(7), 520–534. https://doi.org/10.1038/nrg1110

47. Hansen, A. S., Hsieh, T.-H. S., Cattoglio, C., Pustova, I., Saldaña-Meyer, R., Reinberg, D., Darzacq, X., & Tjian, R. (2019). Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. *Molecular Cell*, *76*(3), 395-411.e13. https://doi.org/10.1016/j.molcel.2019.07.039

48. Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R., & Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *ELife*, *6*, e25776. https://doi.org/10.7554/eLife.25776

49. Hauf, S., Roitinger, E., Koch, B., Dittrich, C. M., Mechtler, K., & Peters, J.-M. (2005). Dissociation of Cohesin from Chromosome Arms and Loss of Arm Cohesion during Early Mitosis Depends on Phosphorylation of SA2. *PLoS Biology*, *3*(3), e69. https://doi.org/10.1371/journal.pbio.0030069

50. Henry, C. A., Urban, M. K., Dill, K. K., Merlie, J. P., Page, M. F., Kimmel, C. B., & Amacher, S. L. (2002). Two linked hairy/Enhancer of split-related zebrafish genes, her1 and her7, function together to refine alternating somite boundaries. *Development*, *129*(15), 3693–3704.

51. Hill, V. K., Kim, J.-S., & Waldman, T. (2016). Cohesin mutations in human cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1866*(1), 1–11. https://doi.org/10.1016/j.bbcan.2016.05.002

52. Hnisz, D., Day, D. S., & Young, R. A. (2016). Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. *Cell*, *167*(5), 1188–1200. https://doi.org/10.1016/j.cell.2016.10.024

53. Horsfield, J. A., Anagnostou, S. H., Hu, J. K.-H., Cho, K. H. Y., Geisler, R., Lieschke, G., Crosier, K. E., & Crosier, P. S. (2007). Cohesin-dependent regulation of Runx genes. *Development*, *134*(14), 2639–2649. https://doi.org/10.1242/dev.002485

54. Horsfield, J., Print, C. G., & Mönnich, M. (2012). Diverse Developmental Disorders from The One Ring: Distinct Molecular Pathways Underlie the Cohesinopathies. *Frontiers in Genetics*, *3*. https://doi.org/10.3389/fgene.2012.00171

55. Hou, F., & Zou, H. (2005). Two Human Orthologues of Eco1/Ctf7 Acetyltransferases Are Both Required for Proper Sister-Chromatid Cohesion. *Molecular Biology of the Cell*, *16*(8), 3908–3918. https://doi.org/10.1091/mbc.e04-12-1063

56. Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., Collins, J. E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J. C., Koch, R., Rauch, G.-J., White, S., … Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498–503. https://doi.org/10.1038/nature12111

57. Huisman, S., Mulder, P. A., Redeker, E., Bader, I., Bisgaard, A.-M., Brooks, A., Cereda, A., Cinca, C., Clark, D., Cormier-Daire, V., Deardorff, M. A., Diderich, K., Elting, M., Essen, A. van, FitzPatrick, D., Gervasini, C., Gillessen-Kaesbach, G., Girisha, K. M., Hilhorst-Hofstee, Y., … Hennekam, R. C. (2017). Phenotypes and genotypes in individuals with SMC1A variants. *American Journal of Medical Genetics Part A*, *173*(8), 2108–2125. https://doi.org/10.1002/ajmg.a.38279

58. Ivanov, M. P., Ladurner, R., Poser, I., Beveridge, R., Rampler, E., Hudecz, O., Novatchkova, M., Hériché, J.-K., Wutz, G., van der Lelij, P., Kreidl, E., Hutchins, J. R., Axelsson-Ekker, H., Ellenberg, J., Hyman, A. A., Mechtler, K., & Peters, J.-M. (2018). The replicative helicase MCM recruits cohesin acetyltransferase ESCO2 to mediate centromeric sister chromatid cohesion. *The EMBO Journal*, *37*(15), e97150. https://doi.org/10.15252/embj.201797150

59. Just, S., Hirth, S., Berger, I. M., Fishman, M. C., & Rottbauer, W. (2016). The mediator complex subunit Med10 regulates heart valve formation in zebrafish by controlling Tbx2b-mediated Has2 expression and cardiac jelly formation. *Biochemical and Biophysical Research Communications*, *477*(4), 581–588. https://doi.org/10.1016/j.bbrc.2016.06.088

60. Kandoth, C., McLellan, M. D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J. F., Wyczalkowski, M. A., Leiserson, M. D. M., Miller, C. A., Welch, J. S., Walter, M. J., Wendl, M. C., Ley, T. J., Wilson, R. K., Raphael, B. J., & Ding, L. (2013). Mutational landscape and significance across 12 major cancer types. *Nature*, *502*(7471), 333–339. https://doi.org/10.1038/nature12634

61. Kaur, M., Mehta, D., Noon, S. E., Deardorff, M. A., Zhang, Z., & Krantz, I. D. (2016). NIPBL expression levels in CdLS probands as a predictor of mutation type and phenotypic severity. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *172*(2), 163–170. https://doi.org/10.1002/ajmg.c.31495

62. Kentepozidou, E., Aitken, S. J., Feig, C., Stefflova, K., Ibarra-Soria, X., Odom, D. T., Roller, M., & Flicek, P. (2020). Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains. *Genome Biology*, *21*(1), 5. https://doi.org/10.1186/s13059-019-1894-x

63. Kim, Y., Shi, Z., Zhang, H., Finkelstein, I. J., & Yu, H. (2019). Human cohesin compacts DNA by loop extrusion. *Science (New York, N.Y.)*, *366*(6471), 1345–1349. https://doi.org/10.1126/science.aaz4475

64. Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K., & Herrmann, B. G. (2017). Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. *Developmental Cell*, *42*(5), 514-526.e7. https://doi.org/10.1016/j.devcel.2017.07.021

65. Kon, A., Shih, L.-Y., Minamino, M., Sanada, M., Shiraishi, Y., Nagata, Y., Yoshida, K., Okuno, Y., Bando, M., Nakato, R., Ishikawa, S., Sato-Otsubo, A., Nagae, G., Nishimoto, A., Haferlach, C., Nowak, D., Sato, Y., Alpermann, T., Nagasaki, M., … Ogawa, S. (2013). Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nature Genetics*, *45*(10), 1232–1237. https://doi.org/10.1038/ng.2731

66. Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G. P., Itoh, T., Watanabe, Y., Shirahige, K., & Uhlmann, F. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature*, *430*(6999), 573–578. https://doi.org/10.1038/nature02742

67. Losada, A. (2014). Cohesin in cancer: Chromosome segregation and beyond. *Nature Reviews Cancer*, *14*(6), 389–393. https://doi.org/10.1038/nrc3743

68. Martin, B. L. (2016). Factors that coordinate mesoderm specification from neuromesodermal progenitors with segmentation during vertebrate axial extension. *Seminars in Cell & Developmental Biology*, *49*, 59–67. https://doi.org/10.1016/j.semcdb.2015.11.014

69. Martin, B. L., & Kimelman, D. (2012). Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. *Developmental Cell*, *22*(1), 223–232. https://doi.org/10.1016/j.devcel.2011.11.001

70. Masciadri, M., Ficcadenti, A., Milani, D., Cogliati, F., Divizia, M. T., Larizza, L., & Russo, S. (2018). Recurrence and Familial Inheritance of Intronic NIPBL Pathogenic Variant Associated With Mild CdLS. *Frontiers in Neurology*, *9*. https://doi.org/10.3389/fneur.2018.00967

71. Mastromina, I., Verrier, L., Storey, K. G., & Dale, J. K. (2017). MYC activity is required for maintenance of the Neuromesodermal Progenitor signalling network and for correct timing of segmentation clock gene oscillations. *BioRxiv*, 212035. https://doi.org/10.1101/212035

72. Mehta, D., Vergano, S. A. S., Deardorff, M., Aggarwal, S., Barot, A., Johnson, D. M., Miller, N. F., Noon, S. E., Kaur, M., Jackson, L., & Krantz, I. D. (2016). Characterization of limb differences in children with Cornelia de Lange Syndrome. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *172*(2), 155–162. https://doi.org/10.1002/ajmg.c.31498

73. Meier, M., Grant, J., Dowdle, A., Thomas, A., Gerton, J., Collas, P., O'Sullivan, J. M., & Horsfield, J. A. (2018). Cohesin facilitates zygotic genome activation in zebrafish. *Development*, *145*(1). https://doi.org/10.1242/dev.156521

74. Merkenschlager, M., & Nora, E. P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. *Annual Review of Genomics and Human Genetics*, *17*(1), 17–43. https://doi.org/10.1146/annurev-genom-083115-022339

75. Mhawech, P., & Saleem, A. (2001). Myelodysplastic syndrome: Review of the cytogenetic and molecular data. *Critical Reviews in Oncology/Hematology*, *40*(3), 229–238. https://doi.org/10.1016/S1040-8428(01)00101-9

76. Mikołajewska, E. (2019). Functional abilities in children with Cornelia de Lange syndrome – pilot study. *Annales Academiae Medicae Silesiensis*, *73*, 199–202.

77. Minor, A., Shinawi, M., Hogue, J. S., Vineyard, M., Hamlin, D. R., Tan, C., Donato, K., Wysinger, L., Botes, S., Das, S., & del Gaudio, D. (2014). Two novel RAD21 mutations in patients with mild Cornelia de Lange syndrome-like presentation and report of the first familial case. *Gene*, *537*(2), 279–284. https://doi.org/10.1016/j.gene.2013.12.045

78. Mir, M., Bickmore, W., Furlong, E. E. M., & Narlikar, G. (2019). Chromatin topology, condensates and gene regulation: Shifting paradigms or just a phase? *Development*, *146*(19). https://doi.org/10.1242/dev.182766

79. Mulder, P. A., Huisman, S. A., Hennekam, R. C., Oliver, C., Balkom, I. D. C. van, & Piening, S. (2017). Behaviour in Cornelia de Lange syndrome: A systematic review. *Developmental Medicine & Child Neurology*, *59*(4), 361–366. https://doi.org/10.1111/dmcn.13361

80. Mullegama, S. V., Klein, S. D., Signer, R. H., Vilain, E., & Martinez-Agosto, J. A. (2019). Mutations in STAG2 cause an X-linked cohesinopathy associated with undergrowth, developmental delay, and dysmorphia: Expanding the phenotype in males. *Molecular Genetics & Genomic Medicine*, *7*(2), e00501. https://doi.org/10.1002/mgg3.501

81. Muto, A., & Schilling, T. F. (2017). Zebrafish as a Model to Study Cohesin and Cohesinopathies. In K. Yokomori & K. Shirahige (Eds.), *Cohesin and Condensin: Methods and Protocols* (pp. 177–196). Springer. https://doi.org/10.1007/978-1-4939-6545-8_11

82. Nakamura, F., Kalb, R. G., & Strittmatter, S. M. (2000). Molecular basis of semaphorin-mediated axon guidance. *Journal of Neurobiology*, *44*(2), 219–229. https://doi.org/10.1002/1097-4695(200008)44:2<219::AID-NEU11>3.0.CO;2-W

83. Nasmyth, K. (2011). Cohesin: A catenase with separate entry and exit gates? *Nature Cell Biology*, *13*(10), 1170–1177. https://doi.org/10.1038/ncb2349

84. Nasmyth, K., & Haering, C. H. (2005). The Structure and Function of Smc and Kleisin Complexes. *Annual Review of Biochemistry*, *74*(1), 595–648. https://doi.org/10.1146/annurev.biochem.74.082803.133219

85. Nasmyth, K., & Haering, C. H. (2009). Cohesin: Its Roles and Mechanisms. *Annual Review of Genetics*, *43*(1), 525–558. https://doi.org/10.1146/annurev-genet-102108-134233

86. Nora, E. P., Goloborodko, A., Valton, A.-L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A., & Bruneau, B. G. (2017). Targeted Degradation of CTCF

Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*, *169*(5), 930-944.e22. https://doi.org/10.1016/j.cell.2017.05.004

87. Ocampo-Hafalla, M., Muñoz, S., Samora, C. P., & Uhlmann, F. (n.d.). Evidence for cohesin sliding along budding yeast chromosomes. *Open Biology*, *6*(6), 150178. https://doi.org/10.1098/rsob.150178

88. Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B. S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A. G., & Merkenschlager, M. (2008). Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. *Cell*, *132*(3), 422–433. https://doi.org/10.1016/j.cell.2008.01.011

89. Piché, J., Van Vliet, P. P., Pucéat, M., & Andelfinger, G. (2019). The expanding phenotypes of cohesinopathies: One ring to rule them all! *Cell Cycle (Georgetown, Tex.)*, *18*(21), 2828–2848. https://doi.org/10.1080/15384101.2019.1658476

90. Prieto, I., Suja, J. A., Pezzi, N., Kremer, L., Martínez-A, C., Rufas, J. S., & Barbero, J. L. (2001). Mammalian STAG3 is a cohesin specific to sister chromatid arms in meiosis I. *Nature Cell Biology*, *3*(8), 761–766. https://doi.org/10.1038/35087082

91. Pugacheva, E. M., Kubo, N., Loukinov, D., Tajmul, M., Kang, S., Kovalchuk, A. L., Strunnikov, A. V., Zentner, G. E., Ren, B., & Lobanenkov, V. V. (2020). CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(4), 2020–2031. https://doi.org/10.1073/pnas.1911708117

92. Rahman, S., Jones, M. J. K., & Jallepalli, P. V. (2015). Cohesin recruits the Esco1 acetyltransferase genome wide to repress transcription and promote cohesion in somatic cells. *Proceedings of the National Academy of Sciences*, *112*(36), 11270–11275. https://doi.org/10.1073/pnas.1505323112

93. Rhodes, J. D. P., Haarhuis, J. H. I., Grimm, J. B., Rowland, B. D., Lavis, L. D., & Nasmyth, K. A. (2017). Cohesin Can Remain Associated with Chromosomes during DNA Replication. *Cell Reports*, *20*(12), 2749–2755. https://doi.org/10.1016/j.celrep.2017.08.092

94. Richterova, J., Huraiova, B., & Gregan, J. (2017). Genome Organization: Cohesin on the Move. *Molecular Cell*, *66*(4), 444–445. https://doi.org/10.1016/j.molcel.2017.05.002

95. Romero-Pérez, L., Surdez, D., Brunet, E., Delattre, O., & Grünewald, T. G. P. (2019). STAG Mutations in Cancer. *Trends in Cancer*, *5*(8), 506–520. https://doi.org/10.1016/j.trecan.2019.07.001

96. Row, R. H., & Kimelman, D. (2009). Bmp inhibition is necessary for post-gastrulation patterning and morphogenesis of the zebrafish tailbud. *Developmental Biology*, *329*(1), 55–63. https://doi.org/10.1016/j.ydbio.2009.02.016

97. Saultz, J. N., & Garzon, R. (2016). Acute Myeloid Leukemia: A Concise Review. *Journal of Clinical Medicine*, *5*(3), 33. https://doi.org/10.3390/jcm5030033

98. Shih, A. H., Abdel-Wahab, O., Patel, J. P., & Levine, R. L. (2012). The role of mutations in epigenetic regulators in myeloid malignancies. *Nature Reviews Cancer*, *12*(9), 599–612. https://doi.org/10.1038/nrc3343

99. Smith, N. L., & Kimelman, D. (2020). Establishing The Body Plan. In *The Zebrafish in Biomedical Research* (pp. 81–88). Elsevier. https://doi.org/10.1016/B978-0-12-812431-4.00007-5

100. Sonoda, E., Matsusaka, T., Morrison, C., Vagnarelli, P., Hoshi, O., Ushiki, T., Nojima, K., Fukagawa, T., Waizenegger, I. C., Peters, J.-M., Earnshaw, W. C., & Takeda, S. (2001). Scc1/Rad21/Mcd1 Is Required for Sister Chromatid Cohesion and Kinetochore Function in Vertebrate Cells. *Developmental Cell*, *1*(6), 759–770. https://doi.org/10.1016/S1534-5807(01)00088-0

101. Stormo, B. M., & Fox, D. T. (2018). Interphase cohesin regulation ensures mitotic fidelity after genome reduplication. *Molecular Biology of the Cell*, *30*(2), 219–227. https://doi.org/10.1091/mbc.E17-10-0582

102. Sun, H., Zhang, J., Xin, S., Jiang, M., Zhang, J., Li, Z., Cao, Q., & Lou, H. (2019). Cul4-Ddb1 ubiquitin ligases facilitate DNA replication-coupled sister chromatid cohesion through regulation of cohesin acetyltransferase Esco2. *PLOS Genetics*, *15*(2), e1007685. https://doi.org/10.1371/journal.pgen.1007685

103. Tahsin, N. (2019). *Cohesin and Regulation of Gene Expression During Zebrafish Development* [Thesis, University of Otago]. https://ourarchive.otago.ac.nz/handle/10523/9188

104. Tomonaga, T., Nagao, K., Kawasaki, Y., Furuya, K., Murakami, A., Morishita, J., Yuasa, T., Sutani, T., Kearsey, S. E., Uhlmann, F., Nasmyth, K., & Yanagida, M. (2000).

Characterization of fission yeast cohesin: Essential anaphase proteolysis of Rad21 phosphorylated in the S phase. *Genes & Development*, *14*(21), 2757–2770. https://doi.org/10.1101/gad.832000

105. Tuuli, M. G., & Odibo, A. O. (2018). 140—Roberts Syndrome. In J. A. Copel, M. E. D'Alton, H. Feltovich, E. Gratacós, D. Krakow, A. O. Odibo, L. D. Platt, & B. Tutschek (Eds.), *Obstetric Imaging: Fetal Diagnosis and Care (Second Edition)* (pp. 577-578.e1). Elsevier. https://doi.org/10.1016/B978-0-323-44548-1.00140-6

106. Viny, A. D., Bowman, R. L., Liu, Y., Lavallée, V.-P., Eisman, S. E., Xiao, W., Durham, B. H., Navitski, A., Park, J., Braunstein, S., Alija, B., Karzai, A., Csete, I. S., Witkin, M., Azizi, E., Baslan, T., Ott, C. J., Pe'er, D., Dekker, J., … Levine, R. L. (2019). Cohesin Members Stag1 and Stag2 Display Distinct Roles in Chromatin Accessibility and Topological Control of HSC Self-Renewal and Differentiation. *Cell Stem Cell*, *25*(5), 682-696.e8. https://doi.org/10.1016/j.stem.2019.08.003

107. Wang, J., & Wynshaw-Boris, A. (2004). The canonical Wnt pathway in early mammalian embryogenesis and stem cell maintenance/differentiation. *Current Opinion in Genetics & Development*, *14*(5), 533–539. https://doi.org/10.1016/j.gde.2004.07.013

108. Ward, A., Hopkins, J., Mckay, M., Murray, S., & Jordan, P. W. (2016). Genetic Interactions Between the Meiosis-Specific Cohesin Components, STAG3, REC8, and RAD21L. *G3: Genes, Genomes, Genetics*, *6*(6), 1713–1724. https://doi.org/10.1534/g3.116.029462

109. Watrin, E., Kaiser, F. J., & Wendt, K. S. (2016). Gene regulation and chromatin organization: Relevance of cohesin mutations to human disease. *Current Opinion in Genetics & Development*, *37*, 59–66. https://doi.org/10.1016/j.gde.2015.12.004

110. Watrin, E., & Peters, J.-M. (2006). Cohesin and DNA damage repair. *Experimental Cell Research*, *312*(14), 2687–2693. https://doi.org/10.1016/j.yexcr.2006.06.024

111. Xing, L., Anbarchian, T., Tsai, J. M., Plant, G. W., & Nusse, R. (2018). Wnt/β-catenin signaling regulates ependymal cell development and adult homeostasis. *Proceedings of the National Academy of Sciences*, *115*(26), E5954–E5962. https://doi.org/10.1073/pnas.1803297115

112. Xiong, B., & Gerton, J. L. (2010). Regulators of the Cohesin Network. *Annual Review of Biochemistry*, *79*(1), 131–153. https://doi.org/10.1146/annurev-biochem-061708-092640

113. Yabe, T., & Takada, S. (2012). Mesogenin causes embryonic mesoderm progenitors to differentiate during development of zebrafish tail somites. *Developmental Biology*, *370*(2), 213–222. https://doi.org/10.1016/j.ydbio.2012.07.029

114. Yoshida, K., Toki, T., Okuno, Y., Kanezaki, R., Shiraishi, Y., Sato-Otsubo, A., Sanada, M., Park, M., Terui, K., Suzuki, H., Kon, A., Nagata, Y., Sato, Y., Wang, R., Shiba, N., Chiba, K., Tanaka, H., Hama, A., Muramatsu, H., … Ogawa, S. (2013). The landscape of somatic mutations in Down syndrome–related myeloid disorders. *Nature Genetics*, *45*(11), 1293–1299. https://doi.org/10.1038/ng.2759

115. Yu, H.-H., & Moens, C. B. (2005). Semaphorin signaling guides cranial neural crest cell migration in zebrafish. *Developmental Biology*, *280*(2), 373–385. https://doi.org/10.1016/j.ydbio.2005.01.029

116. Yuan, B., Neira, J., Pehlivan, D., Santiago-Sim, T., Song, X., Rosenfeld, J., Posey, J. E., Patel, V., Jin, W., Adam, M. P., Baple, E. L., Dean, J., Fong, C.-T., Hickey, S. E., Hudgins, L., Leon, E., Madan-Khetarpal, S., Rawlins, L., Rustad, C. F., … Liu, P. (2019). Clinical exome sequencing reveals locus heterogeneity and phenotypic variability of cohesinopathies. *Genetics in Medicine*, *21*(3), 663–675. https://doi.org/10.1038/s41436-018-0085-6

117. Zhang, H., Emerson, D. J., Gilgenast, T. G., Titus, K. R., Lan, Y., Huang, P., Zhang, D., Wang, H., Keller, C. A., Giardine, B., Hardison, R. C., Phillips-Cremins, J. E., & Blobel, G. A. (2019). Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature*, *576*(7785), 158–162. https://doi.org/10.1038/s41586-019-1778-y

118. Zhang, N., Kuznetsov, S. G., Sharan, S. K., Li, K., Rao, P. H., & Pati, D. (2008). A handcuff model for the cohesin complex. *Journal of Cell Biology*, *183*(6), 1019–1031. https://doi.org/10.1083/jcb.200801157